PCT/NZ03/00109

10/5[redacted]992

Intellectual Property Office of New Zealand
Te Pou Petipati Hanga Hou

REC'D 15 JUL 2003
WIPO PCT

CERTIFICATE

This certificate is issued in support of an application for Patent registration in a country outside New Zealand pursuant to the Patents Act 1953 and the Regulations thereunder.

I hereby certify that annexed is a true copy of the Provisional Specification as filed on 30 May 2002 with an application for Letters Patent number 519330 made by AGRESEARCH;TEAGASC RESEARCH CENTRE;INSTITUT NATIONAL DE LA RECHERCHE AGRONOMIQUE;NATIONAL UNIVERSITY OF IRELAND.

I further certify that pursuant to a claim under Section 24(1) of the Patents Act 1953, a direction was given that the application proceed in the name of DAVIS, George Henry, GALLOWAY, Susan May, GREGAN, Scott Michael, HANRAHAN, James Patrick, JUENGEL, Jennifer Lee, McNATTY, Kenneth Pattrick, MULSANT, Philippe, POWELL, Richard Patrick.

Dated 3 July 2003.

PRIORITY DOCUMENT
SUBMITTED OR TRANSMITTED IN
COMPLIANCE WITH
RULE 17.1(a) OR (b)

Neville Harris

519330

PATENTS FORM NO. 4

Appln Fee: $50.00

James & Wells ref: 31145/14

PATENTS ACT 1953

PROVISIONAL SPECIFICATION

NEW SEQUENCES FOR ALTERING MAMMALIAN OVARIAN FUNCTION AND OVULATION RATE

I/WE AGRESEARCH LIMITED a New Zealand Company of East Street, Ruakura Campus, Hamilton, New Zealand; TEAGASC RESEARCH CENTRE an Irish Company of Athenry, Co, Galway, Ireland; INSTITUT NATIONAL DE LA RECHERCHE AGRONOMIQUE a French Company of Laboratoire de Genetique Cellulaire, Castanet-Tolosan, France AND NATIONAL UNIVERSITY OF IRELAND an Irish Company of Galway, Ireland.

do hereby declare this invention to be described in the following statement:

NEW SEQUENCES FOR ALTERING MAMMALIAN OVARIAN FUNCTION AND OVULATION RATE

TECHNICAL FIELD

The present invention relates to new sequences for altering mammalian ovarian function and ovulation rate.

In particular, the invention broadly concerns a novel mutation in the GDF 9 gene and two novel mutations in the GDF 9B gene. These mutations have been found to be involved in increasing the ovulation rate in heterozygous female mammals; or causing sterility in homozygous female mammals.

BACKGROUND ART

The genes GDF9 and GDF9B (also known as BMP15) code for proteins which are expressed exclusively in the oocyte of the developing follicle, and which play an essential role in mammalian fertility. GDF9 is a member of the transforming growth factor beta (TGFβ) superfamily (McPherron and Lee, 1993) which is expressed in oocytes from the primary stage of follicular development until ovulation (McGrath *et al.*, 1995; Laitinen *et al.*, 1998). GDF9B is closely related to GDF9 (Dube *et al.*, 1998; Laitinen *et al.*, 1998) and is expressed in mouse oocytes at the same time as GDF9, but in human primary follicles slightly later than GDF9. In the ovary GDF9 and GDF9B have now been shown to be expressed exclusively in the developing oocyte in humans (Aaltonen *et al.*, 1999), rodents (Laitinen *et al.*, 1998; Dube *et al.*, 1998; Jaatinen *et al.*, 1999), ruminants (Bodensteiner *et al.*, 1999; Bodensteiner *et al.*, 2000; Galloway *et al.*, 2000) and marsupials (Eckery *et al.*, 2002). In sheep expression of GDF9 can be seen in primordial follicles whereas GDF9B is expressed in primary follicles (Bodensteiner *et al.*, 1999; Galloway *et al.*, 2000).

GDF9 is an essential growth factor for folliculogenesis in mice. Female GDF9 knockout mice (GDF9 -/-) are infertile due to a block in follicular development at the

primary stage (Dong *et al.*, 1996). GDF9B does not appear to be crucial for mouse folliculogenesis as knockout female mice (BMP15 -/-) are fertile (Yan *et al.*, 2001), even though fecundity is somewhat reduced. However, GDF9B is essential for folliculogenesis in sheep as those carrying two copies of naturally-occurring inactivating GDF9B mutations are infertile due to a block in follicular development at the primary stage (Galloway *et al.*, 2000).

In sheep it is also clear that heterozygotes carrying inactivating mutations in one copy of GDF9B (whereby only one copy of the gene produces active protein) have an increased ovulation rate (Galloway *et al.*, 2000). A similar increase in ovulation rate in heterozygote mice with knockouts in either GDF9 or GDF9B has not been observed (Yan *et al.*, 2001). Double knockouts of both GDF9 and GDF9B in mice are infertile with a similar phenotype to GDF9 -/- mice alone, but GDF9B knockout mice (BMP15 -/-) with one active copy of GDF9, have a lower fecundity than BMP15 -/- females (Yan *et al.*, 2001), suggesting that the relative dose of these gene products may also play a role in mice. Collectively these findings suggest that important differences exist in the actions of GDF9 and GDF9B between species with a high ovulation rate phenotype (e.g. mice, rats) and those with a low ovulation rate phenotype (e.g. sheep, humans).

GDF9 maps to a region of sheep chromosome 5 (Sadighi et al., 2002) which is syntenic to the map locations for GDF9 on human chromosome 5 and mouse chromosome 11 (Mouse Genome Informatics (2002). GDF9B maps to the sheep X chromosome (Galloway et al., 2000) in a region of the chromosome syntenic to the map locations for GDF9B on the human and mouse X chromosomes (Dube *et al.*, 1998; Aaltonen *et al.*, 1999). It has also recently been mapped to the pig X chromosome p11-p13 region (Grapes and Rothschild, 2002)

GDF9 and GDF9B, like other members of the TGFβ family, are coded as prepropeptides containing a signal peptide, a proregion and a C-terminal mature region

which is the biologically active peptide. Cleavage of the mature region from the proregion is carried out by an intracellular furin-like protease, and occurs at a conserved furin protease cleavage site. Members of the TGFβ superfamily are biologically active as dimers, and although GDF9 and GDF9B do not contain the cysteine molecule responsible for covalent interchain disulphide bonding seen in other members of the family, these molecules are thought to be biologically active as dimers (Galloway *et al.*, 2000; Yan et al., 2001). However it is unclear whether the physiologically active dimers are homodimers (GDF9-GDF9 and GDF9B-GDF9B), or heterodimers (GDF9-GDF9B) or whether all three dimer forms play a role. It has been postulated based on the above models that GDF9 homodimers play a more important role in the mouse but in sheep the GDF9B homodimers are the most bioactive (Yan *et al.*, 2001). It is unclear whether any such difference is related to the fact that sheep are mono-ovulatory animals (maturing usually only one egg per cycle) whereas mice are poly-ovulatory. Clearly both GDF9 and GDF9B play crucial roles in controlling and maintaining fertility in mammals, and understanding the nature of their actions is essential for the development of therapies.

GDF9 and GDF9B in sheep

The sheep GDF9 gene spans about 2.5 kb and contains 2 exons separated by a single 1126 bp intron (Bodensteiner *et al.*, 1999). The full length coding sequence is 1359 nucleotides long and encodes a pre-propeptide of 453 amino acid residues (Genbank accession number AF078545). A pre-pro region of 318 residues includes a predicted signal sequence, and ends with the RHRR furin protease cleavage site at residues 315 – 318. Residues 319 to 453 beyond the cleavage site code for the 135 amino acid mature active peptide. The amino acid sequence of the sheep GDF9 mature peptide is 92.8 % similar to the human mature peptide and 87.1 % similar to the mouse mature peptide.

Sheep GDF9B has previously been sequenced by us (Galloway *et al.* 2000; Genbank accession nos. AF236078, AF236079) and has a very similar gene structure to GDF9.

The full length coding sequence of 1179 nucleotides is contained in two exons, separated by an intron of about 5.4 kb, and encodes a pre-propeptide of 393 amino acid residues. A pre-pro region of 268 residues includes a predicted signal sequence, and ends with the RRAR furin protease cleavage site at residues 265 – 268. Residues 269 to 393 beyond the cleavage site code for a 125 amino acid mature active peptide. The amino acid sequence of the sheep GDF9B mature peptide is 78.3 % similar to the human mature peptide and 78.6 % similar to the mouse mature peptide.

We have previously shown that the effects on prolificacy in Inverdale and Hanna sheep is due to naturally-occurring mutations in GDF9B (Galloway *et al.*, 2000). Both Inverdale and Hanna sheep have increased ovulation rates in heterozygous carriers of mutated GDF9B, but female homozygous carriers are infertile with 'streak' ovaries (Davis *et al.*, 2001). Infertility in these sheep is due to primary ovarian failure caused by the inability of the follicle to develop beyond the primary stage. Hanna sheep have a single C to T mutation at nucleotide 871 of the GDF9B coding sequence (nucleotide 67 of the mature GDF9B peptide coding region) which produces a premature stop codon in the place of a glutamic acid (Q) at amino acid residue 291 (residue 23 of the mature protein). Inverdale sheep have a distinct T to A mutation at nucleotide 896 (nucleotide 92 of the mature GDF9B peptide coding region) which substitutes valine (V) for aspartic acid (D) at residue 299 (residue 31 of the mature peptide). This substitution of a hydrophobic valine with a negatively charged aspartate changes the electrostatic surface potentials of an area involved in dimer formation and appears to disrupt dimerisation and hence abolish biological activity (Galloway *et al.*, 2000).

In addition to the Inverdale and Hanna lines of sheep discussed above, the Cambridge and F700 Belclare strains of sheep have also been shown to carry genes affecting prolificacy as evidenced in high ovulation rate (Hanrahan & Owen, 1985; Hanrahan, 1991) and the presence of sterile ewes with 'streak-like' ovaries (Hanrahan, 1991; Hanrahan, 1996).

The Cambridge breed was established at the Cambridge University farm in 1964 by screening 54 ewes selected for their high prolificacy from nine British sheep breeds (Owen, 1991). Ewes within the screened flock were subsequently selected on high litter size. Ewes with the highest ovulation rates were selected from this flock in 1984 to provide the foundation animals for the flock now maintained at Teagasc Sheep Research Centre in Ireland (Hanrahan, 1991). A progeny test of 10 Cambridge rams, descended from the flock in Ireland, gave progeny mean ovulation rates ranging from 2.1 - 4.2 (Hanrahan, 1996).

The Belclare breed was established in 1978 at the Belclare Research Centre of Teagasc in Ireland by crossing three populations of prolific sheep assembled by Teagasc in Ireland. These were Fingalway, High Fertility, and Lleyn sheep (Hanrahan, 1989; Hanrahan, 1991). The Fingalway was an interbred cross (from F1) of the Finnish Landrace and Galway breeds; the Lleyn is a breed native to north west Wales and selected animals were imported into Ireland in 1975 by Teagasc for the purposed of developing the Belclare breed; the High Fertility was developed in Ireland during the 1960s from ewes with exceptional litter size performance collected from farms in Ireland between 1963 and 1965. The details of the breed composition of the foundation animals for High Fertility line were given by Hanrahan (1984). A subline of the Belclare (called F700 line) was derived from Belclare sheep that had exceptionally high ovulation rates (Hanrahan 1991). Progeny of 10 Belclare rams had mean ovulation rates ranging from 1.9 – 4.2 (Hanrahan, 1996).

We describe here new naturally-occurring mutations in sheep GDF9 and GDF9B. We show for the first time that mutation of GDF9 causes increased ovulation rate as well as infertility in a manner similar to inactivating mutations in GDF9B, and that GDF9 is also essential for maintaining normal ovarian folliculogenesis in sheep. Furthermore, we show, for the first time in any species, that sheep which are heterozygous for both GDF9 and GDF9B mutations have higher ovulation rates than sheep that are

heterozygous for GDF9 or GDF9B mutations alone ; these observations are supported by genotype, phenotype and immunisation data.

All references, including any patents or patent applications cited in this specification are hereby incorporated by reference. No admission is made that any reference constitutes prior art. The discussion of the references states what their authors assert, and the applicants reserve the right to challenge the accuracy and pertinency of the cited documents. It will be clearly understood that, although a number of prior art publications are referred to herein, this reference does not constitute an admission that any of these documents form part of the common general knowledge in the art, in New Zealand or in any other country.

It is acknowledged that the term 'comprise' may, under varying jurisdictions, be attributed with either an exclusive or an inclusive meaning. For the purpose of this specification, and unless otherwise noted, the term 'comprise' shall have an inclusive meaning - i.e. that it will be taken to mean an inclusion of not only the listed components it directly references, but also other non-specified components or elements. This rationale will also be used when the term 'comprised' or 'comprising' is used in relation to one or more steps in a method or process.

It is an object of the present invention to address the foregoing problems or at least to provide the public with a useful choice.

Further aspects and advantages of the present invention will become apparent from the ensuing description which is given by way of example only.

DISCLOSURE OF INVENTION

The present invention is concerned with novel mutated GDF 9 and GDF 9B gene sequences which alter mammalian ovarian function and ovulation rate. The invention broadly has application in increasing or decreasing the ovulation rate, or causing

sterility in a female mammal, and additionally encompasses regulation of the function of the corpeus luteum.

In particular, the present invention concerns a novel mutation in GDF 9 which increases ovulation rate in heterozygotes and causes sterility in homozygotes for this gene.

The present invention also concerns two mutations in GDF 9B which increase the ovulation rate in heterozygotes for either, but not both together, mutations of the GDF 9B gene. Mammals which are heterozygotes for both mutations in GDF 9B (where each mutation is on a separate X chromosome) are sterile.

The inventors have also discerned that in female mammals that are heterozygous for the mutated GDF 9 gene and heterozygous for one, but not both, of the GDF 9B gene mutations, an even higher ovulation rate exists than in animals heterozygous for one mutation in either GDF 9 or GDF 9B.

Knowledge of the mutated gene sequences can be applied to a test for identifying heterozygous or homozygous female and male mammals carrying the mutated gene. This knowledge of the biological function of the gene and its mutations can also be utilised to increase or decrease the ovulation rate of female mammals, or to induce sterility or reduced fertility in female mammals.

According to a first aspect of the present invention there is provided an isolated mutated GDF 9 nucleic acid molecule comprising a nucleotide sequence selected from the group consisting of:

a) SEQ ID NOs. 1,3 or 5;

b) a sequence complementary to the molecule defined in a);

c) a functional fragment or variant of the sequences in a) or b);

d) an anti-sense sequence to any of the molecules defined in a), b)or c).

Throughout this specification it should understood that the nucleic acid molecule may be a RNA, cRNA, genomic DNA or cDNA molecule, and may be single or double-stranded. The nucleic acid molecule may also optionally comprise one or more synthetic non-natural or altered nucleotide bases, or combinations thereof.

According to a second aspect of the present invention there is provided an isolated mutated GDF 9B nucleic acid molecule comprising a nucleotide sequence selected from the group consisting of:

a) SEQ ID NOs. 7,9,11,13,15 or 17;

b) a sequence complementary to the molecule defined in a)

c) an anti-sense sequence to any of the molecules defined in a) or b).

According to a third aspect of the present invention there is provided an isolated mutated GDF 9 polypeptide comprising an amino acid sequence selected from the group consisting of:

a) SEQ ID NOs. 2,4 or 6;

b) a functional fragment or variant of the sequences in a).

According to a fourth aspect of the present invention there is provided an isolated mutated GDF 9B polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NOs. 8,10,12,14,16 or 18.

According to a fifth aspect of the present invention there is provided an isolated GDF 9 nucleic acid molecule comprising a nucleotide sequence consisting of a mutated nucleotide sequence which encodes a non-conservative amino acid substitution, in the

GDF 9 polypeptide, in at least one codon of the nucleotide sequence associated with receptor binding.

According to a sixth aspect of the present invention there is provided an isolated GDF 9 nucleic acid molecule comprising a nucleotide sequence consisting of a mutated nucleotide sequence which encodes a non-conservative amino acid substitution, in the GDF 9 polypeptide, in at least one codon of the nucleotide sequence associated with dimerisation of the encoded peptide.

According to a seventh aspect of the present invention there is provided an isolated GDF 9B nucleic acid molecule comprising a nucleotide sequence consisting of a mutated nucleotide sequence which encodes a non-conservative amino acid substitution occurring in at least one codon of the nucleotide sequence associated with receptor binding; wherein said nucleotide sequence does not comprise SEQ ID NO. 5 disclosed in WO 01/85926.

According to an eighth aspect of the present invention there is provided an isolated GDF 9B nucleic acid molecule comprising a nucleotide sequence consisting of a mutated nucleotide sequence which encodes a non-conservative amino acid substitution occurring in at least one codon of the nucleotide sequence associated with dimerisation of the peptide; wherein said nucleotide sequence does not comprise SEQ ID NO. 5 disclosed in WO 01/85926.

Suitable programs for ascertaining the structure of polypeptides from the amino acid sequence which can used to determine the regions of the nucleotide sequence associated with dimerisation and/or receptor binding will be known to persons skilled in the art. Examples of suitable computer programs include The Modeller by Rockerfeller University and The SWISS Model developed by Swiss Protein database.

According to a ninth aspect of the present invention there is provided a method of identifying an mammal which carries a mutated nucleic acid molecule encoding GDF-9B, said method comprising the steps of:

i) obtaining a tissue or blood sample from the mammal;

ii) isolating DNA from the sample; and optionally

iii) isolating GDF-9B DNA from the DNA obtained at step i) or ii);

iv) probing said DNA with a probe complementary to either strand of the mutated GDF 9B DNA of SEQ ID NOs 11 or 17;

v) amplifying the amount of mutated GDF 9B DNA;

vi) determining whether the GDF 9B sequence DNA obtained in step v) carries a mutation associated with sterility or increased ovulation.

According to a tenth aspect of the present invention there is provided a method of identifying an mammal which carries a mutated nucleic acid molecule encoding GDF-9, said method comprising the steps of:

i) obtaining a tissue or blood sample from the mammal;

ii) isolating DNA from the sample; and optionally

iii) isolating GDF-9 DNA from the DNA obtained at step i) or ii);

iv) probing said DNA with a probe complementary to either strand of the mutated GDF 9 DNA of SEQ ID NO 5;

v) amplifying the amount of mutated GDF 9 DNA;

vi) determining whether the GDF 9 sequence DNA obtained in step v) carries a mutation associated with sterility or increased ovulation.

According to an eleventh aspect of the present invention there is provided the use of a nucleic acid molecule which is complementary to either strand of the mutated DNA of SEQ ID NOs. 11 or 17 as a marker to identify a mammal carrying a mutated nucleic acid molecule encoding GDF-9B.

The term 'either strand' refers to both the strand of DNA shown in the Sequence ID Number that is being referred to or its complementary strand which is not shown in the sequence listing but which can be determined therefrom.

According to a twelfth aspect of the present invention there is provided the use of a marker as claimed above wherein the identification of mammals carrying said mutated GDF-9B is for DNA assisted selection of mammals that either have enhanced ovulation or sterility.

According to a thirteenth aspect of the present invention there is provided the use of a nucleic acid molecule which is complementary to either strand of the mutated DNA of SEQ ID NO 5 as a marker to identify a mammal carrying a mutated nucleic acid molecule encoding GDF-9.

According to a fourteenth aspect of the present invention there is provided the use of a marker as claimed above wherein the identification of mammals carrying said mutated GDF-9 is for DNA assisted selection of mammals that either have enhanced ovulation or sterility.

According to a fifteenth aspect of the present invention there is provided a probe capable of specifically hybridising to either strand of the mutated GDF 9B DNA of SEQ ID NOs 11 or 17.

According to a sixteenth aspect of the present invention there is provided a probe capable of specifically hybridising to either strand of the mutated GDF 9 DNA of SEQ ID NO 5.

According to a seventeenth aspect of the present invention there is provided a construct comprising a nucleic acid molecule substantially as described above.

According to an eighteenth aspect of the present invention there is provided a vector comprising a nucleic acid molecule substantially as described above.

According to a nineteenth aspect of the present invention there is provided a host cell which comprises a construct or vector substantially as described above which has been introduced therein.

The term "host cell" refers to a cell which is capable of containing a vector or construct and supports the replication and/or expression of the vector or construct. Suitable hosts cells may include *E.coli*, yeast or mammalian cells but should not be limited thereto.

According to a twentieth aspect of the present invention there is provided a cell line comprising a host cell substantially as described above.

According to a twenty-first aspect of the present invention there is provided a method altering the GDF 9 and/or GDF 9B polypeptide composition of a female mammal so as to modulate ovulation comprising the steps of introducing to the genetic material of the mammal at least one nucleic acid molecule comprising a nucleotide sequence selected from the group consisting of:

- SEQ ID NOs 1 or a functional fragment or variant thereof ; and

- SEQ ID NOs 7 or 13 but not both; or

- SEQ ID NOs 3 or a functional fragment or variant thereof; and

- SEQ ID NOs 9 or 15 but not both.

According to a twenty-second aspect of the present invention there is provided the method above wherein the nucleic acid molecule is introduced by a vector or construct.

According to a twenty-fourth aspect of the present invention there is provided a method of modulating the ovulation of a female mammal comprising the steps of:

a) identifying the nucleotide sequences of GDF 9 or GDF 9B carried by the female mammal;

b) administering as appropriate to a mammal depending on the GDF 9 or GDF 9B genes present in the mammal a partial immunisation of a composition comprising:

i) a GDF 9 polypeptide or a functional fragment or variant of GDF9; and/or

ii) a GDF 9B polypeptide or a functional fragment or variant of GDF 9B; and/or

iii) a GDF 9 polypeptide and a GDF 9B polypeptide or functional fragment or variant of GDF 9 or GDF 9B.

together with a pharmaceutically or veterinarily acceptable carrier and/or diluent;

so as to effectively modulate the ovulation of the mammal.

According to a twenty-fifth aspect of the present invention there is provided the method above wherein the carrier is or includes a mild adjuvant.

The term "mild adjuvant" means an adjuvant that induces a moderate immune response. By way of contrast the term "severe adjuvant" means an adjuvant that induces strong immune response.

According to a twenty-six aspect of the present invention there is provided a method of modulating the ovulation rate of a female mammal comprising the steps of:

a) identifying the nucleotide sequences of GDF 9 or GDF 9B carried by the female mammal;

b) administering as appropriate having regard to the GDF 9 and/or GDF 9B genes present in the mammal, an effective amount of an agent selected from the group consisting of:

1) an immunising effective amount of a GDF 9 polypeptide and/or an immunising effective amount of a GDF 9B polypeptide substantially as described above;

2) antisense nucleic acid molecule(s) directed towards nucleic acid(s) encoding:

i) a GDF 9 polypeptide substantially as described above; and/or

ii) a GDF 9B polypeptide substantially as described above.

According to a twenty-seventh aspect of the present invention there is provided a method for breeding a mammal having increased ovulation comprising the steps of:

a) identifying the nucleotide sequences of GDF 9 or GDF 9B carried by the female mammal it is proposed to breed from;

b) identifying the nucleotide sequences of GDF 9 or GDF 9B carried by the male mammal it is proposed to breed from;

c) selecting the female and male animals that will result in progeny having the following characteristics:

i) a single copy of a mutated GDF 9 nucleotide sequence comprising:

A) SEQ ID NO 5; or

B) a functional variant or fragment of the molecule in A) ;or

C) a sequence complementary to the molecule in A) or B); and/or

ii) a single copy of mutated GDF 9B nucleotide sequence comprising:

A) SEQ ID NOs 11 or 17; or

B) a sequence complementary to the molecule(s) in A).

According to a twenty-eighth aspect of the present invention there is provided the method above wherein the mammals selected for breeding will result in progeny having the following characteristics:

i) a single copy of a mutated GDF 9 nucleotide sequence comprising:

A) SEQ ID NO 5; or

B) a functional variant or fragment of the molecule in A) ;or

C) a sequence complementary to the molecule in A) or B);

ii) a single copy of a mutated GDF 9B nucleotide sequence comprising:

A) SEQ ID NOs 11 or 17; or

B) a sequence complementary to the molecule(s) in A).

According to a twenty-ninth aspect of the present invention there is provided a method for selecting a female mammal for breeding on the basis of possessing an increased rate of ovulation comprising the steps of identifying a female mammal possessing only a single mutated copy of:

1) a mutated GDF 9 nucleotide sequence comprising:

a) SEQ ID NO 5; or

b) a functional variant of the molecule of a); or

c) a sequence complementary to the molecules in a) or b);

and/or

2) a mutated GDF 9B nucleotide sequence comprising:

a) SEQ ID NOs 11 or 17; or

b) a sequence complementary to the molecules in a).

According to a thirtieth aspect of the present invention there is provided the method above wherein the mammal selected has both a single mutated copy of GDF 9 and GDF 9B.

According to thirty-first aspect of the present invention there is provided a composition comprising:

i) a mutated GDF 9 polypeptide comprising an amino acid sequence selected from the group consisting of:

A) SEQ ID NOs. 2, 4 or 6; or

B) a functional fragment or variant of the sequences in A); and/or

ii) a mutated GDF 9B polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NOs. 8, 10,12,14,16 or 18

together with a pharmaceutically or veterinarily acceptable carrier and/or diluent.

According to a thirty-second aspect of the present invention there is provided a nucleic acid molecule encoding a polypeptide substantially as described above.

According to a thirty-third aspect of the present invention there is provided a method of modifying the function of the corpus luteum by administering supplementary GDF9 or GDF 9 B , or analogues thereof, or GDF9 or GDF9B antagonists to female mammals.

According to a thirty-fourth aspect of the present invention there is provided a method of modulating the ovulation rate of a female mammal comprising the steps of altering the endogenous levels of GDF 9 and GDF 9B in said mammal.

The present invention also encompasses ligands as polypeptides substantially as described above.

The term "ligand" refers to any molecule which can bind to another molecule such as a polypeptides or peptide, and should be taken to include, but not be limited to, antibodies, cell surface receptors or phage display molecules.

It should be appreciated that the term "antibody" encompasses fragments or analogues of antibodies which retain the ability to bind to a polypeptide of the invention, including but not limited to Fr, $F(ab)_2$ fragments, ScFv molecules and the like. The antibody may be polyclonal but is preferably monoclonal. In some embodiments the ligand may be a phage display molecule.

The term "analogues" above refers to a compound which has a biological function with improved characteristics over the native compounds (e g such a analogue may have a longer half-life than the native compound.)

The term "antagonist" refers to a compound which inhibits the effect of another compound. In this context, the antagonist could refer to a purified antibody, a sera or serum containing an antibody or a plasma or pool of plasma containing an antibody that would neutralise GDF 9 or GDF 9B.

The term "partial immunisation" refers to immunisation of an animal either active or passive of sufficient antigen/antibody to allow for instigation of an immune response to be mounted against the antigen; but the degree of antigen/antibody administered and/or the means of administration are such that insufficient antibodies are produced by the immunised animal to effectively neutralise all the antigen of interest.

The term "full immune response" refers to the immune response of animal which has been fully immunised i.e. the response mounted by the immunised animal results in production of sufficient antibodies to effectively neutralise all the antigen of interest.

The term "protein, or polypeptide" refers to a protein encoded by the nucleic acid molecule of the invention, including fragments, mutations and homologues having the same biological activity i.e. ovulation modulating activity. The protein or polypeptide of the invention can be isolated from a natural source, produced by the expression of a recombinant nucleic acid molecule, or chemically synthesized.

It is to be clearly understood that the invention also encompasses peptide analogues, which include but are not limited to the following:

1. Compounds in which one or more amino acids is replaced by its corresponding D-amino acid. The skilled person will be aware that

retro-inverso amino acid sequences can be synthesised by standard methods; see for example Choreo and Goodman, 1993;

2. Peptidomimetic compounds, in which the peptide bond is replaced by a structure more resistant to metabolic degradation. See for example Olson et al, 1993; and

3. Compounds in which individual amino acids are replaced by analogous structures for example, gem-diaminoalkyl groups or alkylmalonyl groups, with or without modified termini or alkyl, acyl or amine substitutions to modify their charge.

The use of such alternative structures can provide significantly longer half-life in the body, since they are more resistant to breakdown under physiological conditions.

Methods for combinatorial synthesis of peptide analogues and for screening of peptides and peptide analogues are well known in the art (see for example Gallop et al, 1994; Hogan, 1997).

For the purposes of this specification, the term "peptide and peptide analogue" includes compounds made up of units which have an amino and carboxy terminus separated in a 1,2, 1,3, 1,4 or larger substitution pattern. This includes the 20 naturally-occurring or "common" α-amino acids, in either the L or D configuration, the biosynthetically-available or "uncommon" amino acids not usually found in proteins, such as 4-hydroxyproline, 5-hydroxylysine, citrulline and ornithine; synthetically-derived α-amino acids, such as α-methylalanine, norleucine, norvaline, Cα- and *N*-alkylated amino acids, homocysteine, and homoserine; and many others as known in the art.

It also includes compounds that have an amine and carboxyl functional group separated in a 1,3 or larger substitution pattern, such as β-alanine, γ-amino butyric acid, Freidinger lactam (Freidinger *et al*, 1982), the bicyclic dipeptide (BTD)

(Freidinger *et al*, 1982; Nagai and Sato, 1985), amino-methyl benzoic acid (Smythe and von Itzstein, 1994), and others well known in the art. Statine-like isosteres, hydroxyethylene isosteres, reduced amide bond isosteres, thioamide isosteres, urea isosteres, carbamate isosteres, thioether isosteres, vinyl isosteres and other amide bond isosteres known to the art are also useful for the purposes of the invention.

A "common" amino acid is a L-amino acid selected from the group consisting of glycine, leucine, isoleucine, valine, alanine, phenylalanine, tyrosine, tryptophan, aspartate, asparagine, glutamate, glutamine, cysteine, methionine, arginine, lysine, proline, serine, threonine and histidine. These are referred to herein by their conventional three-letter or one-letter abbreviations.

An "uncommon" amino acid includes, but is not restricted to, one selected from the group consisting of D-amino acids, homo-amino acids, N-alkyl amino acids, dehydroamino acids, aromatic amino acids (other than phenylalanine, tyrosine and tryptophan), ortho-, meta- or para-aminobenzoic acid, ornithine, citrulline, norleucine, α-glutamic acid, aminobutyric acid (Abu), and α-α disubstituted amino acids.

The term "introducing" (or grammatical variations thereof) when used in the context of inserting a nucleic acid molecule into a cell, means "transfection" or "transformation" or "transduction" and includes reference to any method for incorporation or transfer of a nucleic acid molecule into a eukaryotic or prokaryotic cell for expression or replication thereof (for example this may include but should not be limited to insertion of a nucleic acid into: a chromosome, mitochondrial DNA, an autonomous replicon (eg. a plasmid). The term "transduction" as used herein, refers to the process of transferring genetic information from a nucleic acid molecule from one cell to another by way of a viral vector.

The term "transfection" as used herein, refers to the uptake, incorporation, and expression of recombinant DNA by eukaryotic cells.

The term "transformation" as used herein refers to a process by which the genetic material carried by an individual cell is altered by incorporation of exogenous DNA into its genome.

The term "variant" as used herein refers to nucleotide and polypeptide sequences wherein the nucleotide or amino acid sequence exhibits substantially 60% or greater homology with the nucleotide or amino acid sequence of the Figures, preferably 75% homology and most preferably 90-95% homology to the sequences of the present invention. – as assessed by GAP or BESTFIT (nucleotides and peptides), or BLASTP (peptides) or BLAST X (nucleotides). The variant may result from modification of the native nucleotide or amino acid sequence by such modifications as insertion, substitution or deletion of one or more nucleotides or amino acids or it may be a naturally-occurring variant. The term "variant" also includes homologous sequences which hybridise to the sequences of the invention under standard hybridisation conditions defined as 2 x SSC at 65°C, or preferably under stringent hybridisation conditions defined as 6 x SCC at 55°C, provided that the variant is capable modulating the ovulation rate of a female mammal or altering ovarian function. Where such a variant is desired, the nucleotide sequence of the native DNA is altered appropriately. This alteration can be effected by synthesis of the DNA or by modification of the native DNA, for example, by site-specific or cassette mutagenesis. Preferably, where portions of cDNA or genomic DNA require sequence modifications, site-specific primer directed mutagenesis is employed, using techniques standard in the art.

A "fragment" of a nucleic acid is a portion of the nucleic acid that is less than full length, and comprises at leàst a minimum sequence capable of hybridizing specifically with a nucleic acid molecule according to the invention, or a sequence complementary thereto, under stringent conditions as defined below. A "fragment" of a polypeptide is a portion of the polypeptide which is less than full length, but which still retains the biological function of either; increasing or decreasing the ovulation rate of a mammal,

causing sterility in a mammal; or altering the regulation of the corpus luteum. Hence, a fragment according to the invention has at least one of the biological activities of the nucleic acid or polypeptide of the invention. However, it will be appreciated that the biological activity of a fragment of the GDF 9 sequence of the present invention encompass only those mutations which will increase the ovulation rate in female mammals heterozygous for the mutation.

The term "isolated" means substantially separated or purified away from contaminating sequences in the cell or organism in which the nucleic acid naturally occurs and includes nucleic acids purified by standard purification techniques as well as nucleic acids prepared by recombinant technology, including PCR technology, and those chemically synthesised.

The term "modulation of ovulation" means increasing or decreasing the rate of ovulation compared to the endogenous rate observed in an untreated animal.

The present invention also includes primers specific for the GDF 9 and GDF 9B nucleotide sequences of the present invention.

The term "hybridization" or grammatical variants thereof means the process of joining two complementary strands of DNA or one each of DNA and RNA to form a double stranded molecule.

"Probes" are single-stranded nucleic acid molecules with a known nucleotide sequence which is labelled in some way (for example, radioactively, fluorescently or immunologically), which are used to find and mark a target DNA or RNA sequence by hybridizing to it. In the present invention the probe will generally be hybridized to the target DNA or RNA sequence under stringent conditions so the probe is specific for the GDF 9 or GDF 9B nucleotide sequences of the present invention.

"Primers" are short nucleic acids, preferably DNA oligonucleotides 15 nucleotides or more in length, which are annealed to a complementary target DNA strand by nucleic acid hybridization to form a hybrid between the primer and the target DNA strand, then extended along the target DNA strand by a polymerase, preferably a DNA polymerase. Primer pairs can be used for amplification of a nucleic acid sequence, e.g. by the polymerase chain reaction (PCR) or other nucleic acid amplification methods well known in the art. PCR-primer pairs can be derived from the sequence of a nucleic acid according to the present invention, for example, by using computer programs intended for that purpose such as Primer (Version 0.5© 1991, Whitehead Institute for Biomedical Research, Cambridge, MA).

Methods for preparing and using probes and primers are described, for example, in Sambrook et al. Molecular Cloning: A Laboratory Manual, 2nd ed, vol. 1-3, ed Sambrook et al. Cold Spring Harbour Laboratory Press, Cold Spring Harbour, NY, 1989.

Probes or primers can be free in solution or covalently or noncovalently attached to a solid support by standard means.

"Stringent conditions" for the amplification of a target nucleic acid sequence (eg by PCR) using a particular amplification primer pair, are conditions that permit the primer pair to hybridize only to the target nucleic acid sequence to which a primer having the corresponding wild type sequence (or its complement) would bind.

Nucleic acid hybridization is affected by such conditions as salt concentration, temperature, or organic solvents, in addition to the base composition, length of the complementary strands, and the number of nucleotide base mismatches between the hybridizing nucleic acids, as will be readily appreciated by those skilled in the art.

When referring to a probe or primer, the term "specific for (a target sequence)" indicates that the probe or primer hybridizes under stringent conditions only to the target sequence in a given sample comprising the target sequence.

The term "construct" as used herein refers to an artificially assembled or isolated nucleic acid molecule which includes the gene of interest. In general a construct may include the gene or genes of interest and appropriate regulatory sequences. It should be appreciated that the inclusion of regulatory sequences in a construct is optional for example, such sequences may not be required in situations where the regulatory sequences of a host cell are to be used. The term construct includes vectors but should not be seen as being limited thereto.

The term "vector" as used herein encompasses both cloning and expression vectors. Vectors are often recombinant molecules containing nucleic acid molecules from several sources.

A "cloning vector" refers to a nucleic acid molecule originating or derived from a virus, a plasmid or a cell of a higher organism into which another exogenous (foreign) nucleic acid molecule of interest, of appropriate size can be integrated without loss of the vector's capacity for self-replication. Thus vectors can be used to introduce at least one foreign nucleic acid molecule of interest (e.g. gene of interest) into host cells, where the gene can be reproduced in large quantities.

An "expression vector" refers to a cloning vector which also contains the necessary regulatory sequences to allow for transcription and translation of the integrated gene of interest, so that the gene product of the gene can be expressed.

The cloning vector may be selected according to the host or host cell to be used. Useful vectors will generally have the following characteristics:

(a) the ability to self-replicate;

(b) the possession of a single target for any particular restriction endonuclease; and

(c) desirably, carry genes for a readily selectable marker such as antibiotic resistance.

Two major types of vector possessing these characteristics are plasmids and bacterial viruses (bacteriophages or phages). Presently preferred vectors are bacterial, insect or mammalian vectors and may include the following: the pUC, pBlueScript, pGEM, PGEX, pBK-CMV, lambda ZAP, lambda GEM, pEFIRES-P, pUB6/V5/His, pBC1, pADTrack-CMV, pAdenovator, pAdEasy-1, pSFV-PD, pCA3, pBABE; pPIC9, pA0815, pET and pSP series. However, this list should not be seen as limiting the scope of the present invention.

Examples of preferred expression systems are as follows:

1. For an *in vitro* cell expression system, the 293T cell system with a pEFIRES-P vector (Hobbs S *et al.*,1998) which confers puromycin resistance may be used. For coexpression of two genes, the aforementioned vector may be modified to change the antibiotic resistance gene to bleomycin resistance. Alternatively, the co-expression of two genes and the selection gene can be achieved by constructing a tricistronic expression vector. A corresponding stably transfected insect cell system can also be used, e.g. the S2 cell system using "DES" vector expression system; www.invitrogen.com.

2. With respect to expressing GDF's in all tissues of transgenic animals, one approach is to use the pUB6/V5-His A vector (www.invitrogen.com) to make the constructs. For tissue-specific expression the rat PEPCK 0.6 kb promoter for liver and kidney expression can be included in the construct by replacing

the Ubi-C promoter in the pUB6/V5-His A vector with the PEPCK promoter. For GDF expression in mammary tissue another promoter system would be preferred. For this tissue one approach would be to use the bovine β-lactoglobulin gene promoter and/or the bovine α S1 casein promoter (e.g. pBC1 vector, www.invitrogen.com) to drive the expression of the GDFs into milk. For global over-expression in transgenic animals, the CMV enhanced β-actin promoter (Okabe M, et al.; FEBS Letters 407: 313-319, 1997) or a modified EF1 α-promoter can be used also (Taboit-Dameron F, et al., Transgenic Research 8: 223-235, 1998).

Adenoviruses, retroviruses and alphaviruses are other suitable mammalian expression systems. A typical approach to those skilled in the art is that described by (TC He *et al.*, 1998),. With respect to GDF expression the pAd Track-CMV vector or pAdenovator vectors (www.qbiogene.com) can be used to make the construct which is then co-transformed with pAd Easy-1 adenoviral plasmid into *E. coli* to generate a recombinant adenoviral genome which contains a CMV-promoter driven GDF expression cassette. This recombinant adenoviral genome is then transfected into 293T cells to make the virus stock. Alternative methods for generating adenoviruses can also be used for the same purpose (e.g. PCA3 plasmid based gene transfer (www.microbix.com); or COS-TPC method (Miyake S *et al.*,1996).

3. Non-cytopathogenic Semliki Forest viruses expressing GDF's can be generated using, for example, pSFV-PD vectors as described by Lundstrom et al., Histochem Cell Biol 115: 83-91, 2001. Furthermore, retroviral expression systems based on, for example, pBABE vectors, can be used for expressing

GDF's in mammalian cells (Morgenstern, JP and Land, H, 1990; Nucleic Acids Res 18: 3587-3596).

4. Yeast cells (e.g. *Pichia pastoris*, *Saccharomyces cerevisiae*) are another well established expression system to those skilled in the art (C Hadfield, *et al.*, 1993);(MA Romanos *et al.*, 1992),. For example, the pPIC9 vector (www.invitrogen.com) can be used in *Pichia pastoris* for the expression of GDF's. For coexpression of two genes, the vector pA0815 (www.invitrogen.com) is a preferred candidate.

5. *Echerichia coli* (*E. coli*) is a standard laboratory expression system in widespread use. For example, the pET expression system (www.novagen.com) can be used to express recombinant mammalian GDF-9 and GDF-9B

The DNA molecules of the invention may be expressed by placing them in operable linkage with suitable control sequences in a replicable expression vector. Control sequences may include origins of replication, a promoter, enhancer and transcriptional terminator sequences amongst others. The selection of the control sequence to be included in the expression vector is dependent on the type of host or host cell intended to be used for expressing the DNA as would be understood by a person skilled in the art.

The term "operably linked" or grammatical variant thereof as used herein means that the regulatory sequences necessary for expression of the gene of interest are placed in the nucleic acid molecule in the appropriate positions relative to the gene to enable expression of the gene.

As used herein the term "regulatory sequences" refers to certain nucleic acid sequences such as origins of replication, promoters, enhancers, polyadenylation signals, terminators and the like, that enable expression of the nucleic acid molecule of interest.

The term "expression" as used herein broadly refers to the process by which a nucleic acid molecule is converted by transcription and then translation into a protein.

The term "gene" as used herein refers to a nucleic acid molecule comprising an ordered series of nucleotides that encodes a gene product (i.e. specific protein).

The expression vectors useful in the present invention may contain at least one expression control sequence that is operatively linked to the DNA sequence or fragment to be expressed. The control sequence is inserted in the vector in order to control and to regulate the expression of the cloned DNA sequence. Examples of useful expression control sequences are the lac system, the trp system, the tac system, the trc system, major operator and promoter regions of phage lambda, the glycolytic promoters of yeast acid phosphatase, e.g. Pho5, the promoters of the yeast alpha-mating factors, and promoters derived from polyoma, adenovirus, retrovirus, and simian virus, and cytomegalovirus e.g. the early and late promoters of SV40, and other sequences known to control the expression of genes of prokaryotic and eukaryotic cells and their viruses or combinations thereof.

In the construction of a vector it is also an advantage to be able to identify the bacterial clone carrying the vector incorporating the foreign DNA. Such assays include measurable colour changes, antibiotic resistance and the like. In one preferred vector, the β-galactosidase gene is used, which gene is detectable by clones exhibiting a blue phenotype on X-gal plates. This facilitates selection. Once selected, the vectors may be isolated from the culture using standard procedures.

Depending on the host used, transformation and transfection is performed according to standard techniques appropriate to such cells. For prokaryotes or other cells that contain substantial cell walls, the calcium treatment process (Cohen, S N Proceedings, National Academy of Science, USA 69 2110 (1972)) may be employed. For mammalian cells without such cell walls, the calcium phosphate precipitation method of Graeme and Van Der Eb, Virology 52:546 (1978) or liposomal reagents are preferred.

Upon transformation of the selected host with an appropriate vector the polypeptide encoded can be produced, often in the form of a fusion protein, by culturing the host cells. The polypeptide of the invention may be detected by rapid assays as indicated above. The polypeptide is then recovered and purified as necessary. Recovery and purification can be achieved using any procedures known in the art, for example by absorption onto and elution from an anion exchange resin. This method of producing a polypeptide of the invention constitutes a further aspect of the present invention.

The preparation of pharmaceutical compositions including pharmaceutical carriers are well known in the art, and are set out in textbooks such as Remington's Pharmaceutical Sciences, 19th Edition, Mack Publishing Company, Easton, Pennsylvania, USA.

The compounds and compositions of the invention may be administered by any suitable route, and the person skilled in the art will readily be able to determine the most suitable route and dose for the condition to be treated. Dosage will be at the discretion of the attendant physician or veterinarian, and will depend on the nature and state of the condition to be treated, the age and general state of health of the subject to be treated, the route of administration, and any previous treatment which may have been administered.

The carrier or diluent, and other excipients, will depend on the route of administration, and again the person skilled in the art will readily be able to determine the most suitable formulation for each particular case.

The invention also includes adenovirus-based gene therapy techniques for expressing GDF-9B and GDF-9/GDF-9B in cell cultures, organ cultures and whole experimental animals for manipulating ovarian follicular protein synthesis or production.

BRIEF DESCRIPTION OF DRAWINGS

Further aspects of the present invention will become apparent from the following description which is given by way of example only and with reference to the accompanying drawings in which:

Figure 1 Shows Irish Cambridge and F700 Belclare sheep GDF9 sequence and mutations;

Figure 2 Shows Irish Cambridge and F700 Belclare sheep GDF9B sequence and mutations;

Figure 3 Shows schematic representation of genotypes within a F700 Belclare and Cambridge pedigrees;

Figure 4 Shows nucleotide and amino acid of wildtype sheep GDF9 showing positions of mutations in Irish Cambridge and F700 Belclare sheep;

Figure 5 Shows nucleotide and amino acid of sheep GDF9B showing positions of mutations in Irish Cambridge and F700 Belclare sheep;

Figure 6 Shows alignment of GDF9 and GDF9B protein sequence with other members of the TGFβ superfamily members for which structures have been determined;

Figure 7 Shows examples of the pattern of progesterone concentrations in plasma of actively immunized ewes; and

Figure 8 Shows the average concentrations of progesterone in plasma following synchronization of luteal regression.

BEST MODES FOR CARRYING OUT THE INVENTION

Non-limiting examples illustrating the invention will now be provided. It will be appreciated that the above description is provided by way of example only and variations in both the materials and techniques used which are known to those persons skilled in the art are contemplated.

METHODOLOGY

Animals

The flocks of Cambridge and Belclare sheep at The Sheep Research Centre of Teagasc, Athenry, are routinely examined for ovulation rate at the beginning of each mating season using laparoscopy. The examination is done once before joining and once after the first mating of the joining period. These data have been collected each year since these flocks were established. The flocks are self-contained with at least 5 males used for mating each year. In addition rams from these flocks have been progeny tested for ovulation rate by crossing with Galway and Scottish Blackface ewes – both low prolificacy breeds. Ovulation rate measurements are done by laparoscopy under licence from The Minister of Health under the Cruelty to Animals Act (1876) EU Directive 86/609/EC.

When sterile ewes were first detected they were checked for the possibility that they were freemartins but this could not be confirmed (Hanrahan, 1991). Blood samples were retained for DNA extraction from the sterile Cambridge ewes born in 1990 and later years and from essentially all of the F700 Belclare sterile females born since

1993. This material has been supplemented by blood samples for DNA extraction collected from fertile ewes in these flock from 1992 onwards.

Ovulation rate data were analysed by least squares procedures with the individual animal as the experimental unit using the GLM procedure of SAS. The factors in the models were ewe, age, year of record, and the number of copies (0 or 1) of each of the mutations described below.

Genomic DNA was isolated from Irish Cambridge and F700 Belclare sheep either from frozen stored buffy coat or directly from white blood cells in whole blood using the method of Montgomery and Sise (Montgomery and Sise, 1990). Parentage of key pedigrees was verified with autosomal sheep microsatellite markers OarHH64 (sheep chromosome 4), OarCP34 (sheep chromosome 3) and OarFCB304 (sheep chromosome 19) (Maddox *et al.*, 2001).

Sequencing and Mutation Detection

The sheep GDF9 and GDF9B genes were amplified using the polymerase chain reaction (PCR) with primers designed from published sheep sequences (sheep genomic GDF9B exon 1, AF236078; sheep genomic GDF9B exon 2, AF236079; sheep genomic GDF9 exon 1 and 2, AF078545).

The PCR primers used were as follows:

GDF9B exon 1	B13: 5'-ACTGCTGCCTTGTCCCAC-3'
	B28: 5'-AGGCAATGTGAAGCCTGACA-3'
GDF9B exon 2	B25: 5'-CAGTTTGTACTGAGCAGGTC-3'
	O4: 5'-TTCTTGGGAAACCTGAGCTAGC-3'
GDF9 exon 1	G1: 5'-GAATTGAACCTAGCCCACCCAC-3'

G4: 5'-AGCCTACATCAACCCATGAGGC-3'

GDF9 exon 2 G5: 5'-ATCCCACCCTGACGTTTAAGGC-3'

G7: 5'-TCCTCCCAAAGGCATAGACAGG-3'

The resulting PCR products were sequenced on an ABI 373 sequencer.

Single Stranded Conformational Polymorphism Detection

SSCP (single stranded conformational polymorphism) was carried out on 9 Belclare rams involved in the progeny testing programme and on the half sib progeny of three of these rams (n = 58 (29, 17, and 12 progeny respectively)) and also on 2 Cambridge rams one of whom was progeny tested. In addition, seven purebred daughters of two of the Belclare rams were tested along with four of the five dams involved.

GDF9B genotypes were determined by analysis of three nucleotide fragments which spanned most of exon 2. Fragments analysed by SSCP were:

353 bp, exon 2	primer 9B-359 5'-CGC TTT GCT CTT GTT CCC TCT-3'
	primer 9B-691 5'-CCT CAC TAC CTC TTG GCT GCT-3'
273 bp, exon 2	primer 9B-664 5'-GGG TTC TAC GAC TCC GCT TC-3'
	primer 9B-916 5'-GGT TAC TTT CAG GCC CAT CAT-3'
312 bp, exon 2	primer 9B-915 5'-CAT GAT GGG CCT GAA AGT AAC-3'
	primer 9B-1205 5'-GGC AAT CAT ACC CTC ATA CTC C-3'

Primers were designed from nucleotide sequence Genbank Accession number AF236079 and primer names correspond to nucleotide position within that sequence.

GDF9 genotypes were determined by analysis of five fragments which spanned exon 1, part of the intron and most of exon 2. Fragments analysed by SSCP were:

462 bp, exon 1	primer G9-1734 5'-GAA GAC TGG TAT GGG GAA ATG-3'
	primer G9-2175 5'-CCA ATC TGC TCC TAC ACA CCT-3'
294 bp, intron	primer G9-2676 5'-GTG TGA GAG AGA TGG GAG CA-3'
	primer G9-2947 5'-AAG AGG AAA ACT ATC AAA AGA CA-3'
296 bp, exon 2	primer G9-3270: 5'-TGG CAT TAC TGT TGG ATT GTT TT-3'
	primer G9-3546: 5'-CAA GAG GAG CCG TCA CAT CA-3'
206 bp, exon 2	primer G9-3543: 5'-GAT TGA TGT GAC GGC TCC TCT-3'
	primer G9-3728: 5'-GGG AAT GCC ACC TGT GAA AAG-3'
221 bp, exon 2	primer G9-3939: 5'-TCT TTT TCC CCA GAA TGA ATG T-3'
	primer G9-4140: 5'-CAC AGG ATG GTC TTG GCA CT-3'

Primers were designed from nucleotide sequence Genbank Accession number AF078545 and primer names correspond to nucleotide position within that sequence.

Amplification was carried out for 30 cycles in a 40 µL reaction mixture, using 150 ng of genomic DNA, with 1.5 mM or 3 mM magnesium and an annealing temperature of 55 to 58° C. PCR fragments were analysed by SSCP in polyacrylamide gels with overnight migration at 9-15 V/cm, 15°C.

Single Nucleotide Polymorphism Detection Assays

The [E1] polymorphism identified in GDF9 exon 1 produced a G to A nucleotide change which disrupts a *Hha* I restriction enzyme cleavage site (GCGC to GCAC) at nucleotide 260 of the 462 bp PCR fragment produced by primers G9-1734 and G9-2175 above. Digestion was carried out using 9 µl of PCR product and 3 U *Hha* I in 15 µl final volume, for 6 h at 37° C. Restriction digestion of the PCR product from wildtype animals with *Hha* I resulted in cleavage of the 462 bp product (at two internal *Hha* I sites) into fragments of 52 bp, 156 bp and 254 bp. However, DNA fragments containing the A nucleotide are not cleaved at this site and fragment sizes of 52 bp and 410 bp are seen. Animals heterozygous for the mutation have fragments of all four sizes (52 bp, 156 bp, 254 bp and 410 bp).

The remaining single nucleotide polymorphisms (SNPs) in GDF9 and GDF9B identified by sequencing did not affect common restriction endonuclease cleavage sites. In order to screen these polymorphisms through the F700 Belclare and Cambridge flocks of sheep, PCR was carried out using primers with single mismatches in order to deliberately generate products that contained restriction enzyme sites. Assays have been designed so that digestion with the appropriate restriction enzyme cleaves either PCR products from wild-type animals or PCR products from animals containing the SNP, as specified below. The resulting band shift can be resolved on a high percentage agarose gel. The primer sequences and PCR conditions for each assay are as follows. The mismatch created in the appropriate primer to generate the restriction enzyme cleavage site is underlined.

In all five assays below, amplification was carried out at: 94°C for 5 min; 35 cycles of 94°C for 30 sec, an annealing step for 40 sec (at the specific temperature stated below for each assay) and 72°C for 30 sec; followed by a final extension of 72°C for 4 min. Magnesium concentration was 1.5 mM.

The primers used for the GDF9 [324] nucleotide change amplify a 161 bp PCR product.

[324]-Sfu1F 5'-GGAATATTCACATGTCTGTAAATTTTACATGTTCG-3'

[324]-Sfu3R 5'-GAGGGAATGCCACCTGTGAAAAGCC-3'

Annealing temperature 63°C

Non-wildtype strand cleaved by restriction enzyme *Sfu* I

The primers used for the GDF9 [714] nucleotide change amplify a 158 bp PCR product.

[714]-Tru1R 5'-CAGTATCGAGGGTTGTATTTGTGTGGGGCCT-3'

[714]-Tru3F 5'-GCCTCTGGTTCCAGCTTCAGTC-3'

Annealing temperature 63°C

Non-wildtype strand cleaved by restriction enzyme *Mse* I

The primers used for the GDF9 [787] nucleotide change amplify a 139 bp PCR product.

[787]-Dde1R: 5'-CATGGATGATGTTCTGCACCATGGTGTGAACCTGA-3'

[787]-Dde3F: 5'-CTTTAGTCAGCTGAAGTGGGACAAC-3'

Annealing temperature 62°C

Wildtype strand cleaved by restriction enzyme *Dde* I

The primers used for the GDF9B [S1] nucleotide change amplify a 141 bp PCR product.

[S1]-Hinf1F: 5'-CACTGTCTTCTTGTTACTGTATTTCAATGAGAC-3'

B26: 5'-GATGCAATACTGCCTGCTTG-3'

Annealing temperature 63°C

Wildtype strand cleaved by restriction enzyme *Hinf* I

The primers used for the GDF9B [S2] nucleotide change amplify a 153 bp PCR product.

[S2]-Dde1F: 5'-GCCTTCCTGTGTCCCTTATAAGTATGTTCCCCTTA-3'

O4: 5'-TTCTTGGGAAACCTGAGCTAGC-3'

Annealing temperature 64°C

Wildtype strand cleaved by restriction enzyme *Dde* I

Restriction digestion of PCR with *Hinf* I [S1] or *Dde* I [[787] and [S2]] resulted in a cleavage of the longer primer from the fragment amplified from wild-type alleles (thus producing a 30-35 bp smaller product than the uncleaved fragment from animals containing alleles with the mutation). Restriction digestion of PCR with *Sfu* I [324] or *Mse* I [714] resulted in a cleavage of the longer primer from the fragment amplified from mutant alleles (thus producing a 30-35 bp smaller product than the uncleaved fragment from animals containing wild-type alleles). Animals heterozygous for any of the mutations have fragments of both sizes. The digested fragments were separated on a 4% agarose gel and visualised with ethidium bromide staining. The gels were scored for the presence or absence of the mutations. Homozygous, heterozygous and negative controls were included with each assay.

Immunisation experiments

All experiments were performed with the approval of the Animals Ethics Committee at Wallaceville Animal Research Centre in accordance with the 1987 Animal Protection (Codes of Ethical Conduct) Regulations of New Zealand. The animals used in the immunization studies (n=54) were 5 to 6 year old parous Romney ewes.

Generation of antigens for immunization of sheep

Peptides KKPLVPASVNLSEYFC (GDF9) and SEVPGPSREHDGPESC (GDF9B) were synthesized and conjugated to KLH through the C terminal cysteine residue by Macromolecular Resources (Colorado State University, Fort Collins, CO).

Active immunization of ewes against GDF9 and GDF9B peptides

Ewes were injected (i.m) with 0.4 mg KLH (control, n=10), 0.4 mg KLH-GDF9 peptide conjugate (GDF9 peptide; n=10) or 0.4 mg KLH-GDF9B peptide conjugate (GDF9B peptide; n=10) in 1 ml of Freund's complete adjuvant for the initial immunization. Thereafter, ewes were immunized once monthly with 0.2 mg KLH, GDF9 peptide or GDF9B peptide in 1 ml of saline mixed with 1.25 ml STM (Span-Tween-Marcol) for 6 months. After the 5th injection, vasectomised rams with marking harnesses were run with the ewes to monitor estrous cycles. The length of the estrous cycle was calculated as the days between first observed markings by the vasectomised ram of successive cycles. In addition, blood samples were collected via the jugular vein 3 times a week for determination of plasma progesterone concentrations. Ovulation rates of the ewes that displayed estrous behaviour were determined by laparoscopy once all of the control ewes had been observed in estrus and for each successive estrous cycle. In addition, ovulation rate of all ewes was determined by laparoscopy 3-4 weeks prior to ovarian collection. Approximately 2 weeks following the final injection, ewes were killed using a captive bolt and exsanguinated. The blood collected from all ewes was to be used in subsequent passive immunisation studies.

Both ovaries were recovered and the number of corpora lutea present was recorded and one ovary from each ewe was fixed in Bouins fluid for morphological examination and analysis of follicular populations.

Passive immunisation of ewes against KLH, KLH-GDF9 peptide and KLH-GDF9B peptide

Pools of antiplasma from KLH (n=9), GDF9 peptide (n=7, all anovulatory ewes) and GDF9B peptide (n=9, all anovulatory ewes) treated ewes were generated by combining the plasmas obtained from some of the actively immunized ewes within each treatment group. The estrous cycles of ewes were synchronized by using a prostaglandin $F_{2\alpha}$ derivative (Estrumate; 125µg). Estrus was detected with the aid of a vasectomised ram wearing a marking harness. On day 4 or 5 of the estrus cycle (estrus = day 0) ewes were laparascoped to determine ovulation rate and fitted with an indwelling jugular cannula. The following day ewes (n=4-5 per group) were administered 100 ml of antiplasma to KLH, GDF9 peptide or GDF9B peptide through the indwelling jugular cannula. Ewes were given another injection of Estrumate, at 96h after administration of the antiplasma to induce a follicular phase and ovulation rate was determine by laparoscopy at 10 days after the injection of Estrumate and every 15-18 days thereafter until the end of the breeding season (as assessed by lack of estrous activity in non-experimental sheep). Blood samples were collected from the ewes at 5 minutes, 1 h and 96 h after injection of the antiplasma and thereafter 3 times a week from the 2nd injection of Estrumate for determination of antibody titers and concentrations of progesterone in plasma.

Determination of progesterone concentrations

Concentrations of progesterone in plasma were determined by RIA as described previously (McNatty et al., 1981). The inter- and intra-assay co-efficients of variation

were <10% and assay sensitivity was 0.1 ng/ml. All samples below the sensitivity of the assay were assigned a value of 0.1 ng/ml for statistical analysis.

Short-term immunisations

Romney ewes were immunised with either KLH, KLH conjugated to GDF9 peptide or KLH conjugated to GDF9B peptide. The antigens were administered in DEAE Dextran (4% w/v) on 2 occasions one month apart. Thereafter the ovaries of these animals were visualised following exteriorisation via a mid-line incision and the number of corpora lutea counted (two observations at successive cycles per ewe).

Statistical analysis

For the long-term, actively immunized ewes, ovulation rate for individual ewes was calculated as the mean of the number of corpora lutea observed at all observations for that ewe when at least 1 corpus luteum (CL) was present (i.e. observations of no CL were excluded from the calculation). The Kruskal-Wallis test was used to compare ovulation rates between the KLH-GDF9B mature protein and the KLH treated groups. No other groups were included in this comparison since none had sufficient numbers of ewes ovulating. The Chi Square test was used to compare the proportion of ewes observed in estrus by the time all the control ewes had been observed in estrus. In addition the Chi Square test was used to compare the proportion of ewes with corpora lutea on their ovaries 3-4 weeks before and at ovarian collection.

When examining the effects of active immunization treatments on ovarian volumes, numbers of follicles or oocyte or follicular diameters, the data were analysed within each follicle type after normalising the data by log transformation. For each parameter a one-way ANOVA was performed, after blocking on animals where appropriate, and differences between treatment groups were determined by least significant difference.

For the passively immunized ewes, differences in the number of ewes with corpora lutea at each laparoscopy were determined using Fisher's exact test. The areas under the curves were calculated using Genstat using the area function for progesterone values from 2 to 19 days following injection of Estrumate that was given 4 days after administration of plasma. Resulting values were analysed with one-way ANOVA and differences between the control and treated ewes determined with Fisher's pairwise comparisons.

For the short term active immunizations, ovulation rate for individual ewes was calculated as the average of the number of corpora lutea observed at both observations. Data was analysed using the general linear models procedures of SAS. Differences between least-squares means were evaluated by least significant differences.

RESULTS and ANALYSIS OF RESULTS

Finding mutations in Cambridge and F700 Belclare animals

In order to determine whether mutations in GDF9 or GDF9B were contributing to sterility in these animals sequence information was obtained for the entire coding sequence of both genes in a subset of Irish Cambridge (N = 9) and F700 Belclare sheep (N = 10). Animals were chosen for full-length sequencing based on their sterility phenotype or their pedigree relationship to sterile animals. In addition, mutation detection was also carried out by single-stranded conformational polymorphism (SSCP) analysis independently of the above sequencing in F700 Belclare pedigrees (23 animals and 58 progeny test daughters of three rams) and also on 2 Cambridge rams.

Mutations in GDF9

Sequence of GDF9 revealed eight single nucleotide polymorphisms across the entire coding region (Table 1, Figure 4). SSCP analysis identified five fragments across the

gene which contained conformational differences. These differences correspond to one single nucleotide polymorphism (SNP) in exon 1, one SNP in the intron and five SNPs in exon 2.

Original naming of the mutations (numbers in square brackets [ ], Table 1) refers to the nucleotide position from the start of exon 2, except for [E1] which refers to the polymorphism found in exon 1 of GDF9. Table 1 shows the relationship between (a) the original numbering system, (b) the coding nucleotide position in the full length coding sequence numbered from the first atg, (c) the position of the coding amino acid residue involved, and (d) the position of the coding residue (if any) within the mature coding sequence, numbering from the first amino acid after the furin protease processing site. For example mutation [581] is a G to A nucleotide substitution at coding nucleotide 978 of GDF9 which corresponds to an unchanged glutamate (Glu) coding residue 326 of the full length unprocessed protein, or residue 8 of the processed mature peptide.

Three of the eight polymorphisms are nucleotide changes which do not result in an altered amino acid ([74] at nucleotide position 471, [80] at nucleotide 477, and [581] at nucleotide position 978). The five remaining nucleotide changes [324], [597], [714], and [787] give rise to amino acid changes (Table 1), Figure 1, Figure 4), although three of them are relatively conservative changes. The [E1] arginine to histidine change at amino acid residue 87 in exon 1 substitutes one basic charged polar group with another, and occurs at a position prior to the furin processing site, so is unlikely to affect the activity of the mature protein. Both the [597] valine to isoleucine change at amino acid residue 332 of the unprocessed protein (residue 14 of the mature coding region) and the [714] valine to methionine at residue 371 of the unprocessed protein (residue 53 of the mature coding region) substitute non-polar groups with non-polar groups. The remaining two changes result in non-conservative substitutions. The [324] glutamic acid to lysine change at amino acid residue 241 of the unprocessed protein

replaces an acidic group with a basic group, but this occurs at a position prior to the furin processing site and is unlikely to affect the mature active coding region. However the [787] serine to phenylalanine change at residue 395 replaces an uncharged polar group with a non-polar group at residue 77 of the mature coding region. The nucleotide and amino acid changes are illustrated in Figure 1 and Figure 4.

Mutations in GDF9B

Both DNA sequencing and independent SSCP analysis of GDF9B in Cambridge and F700 Belclare sheep revealed four polymorphisms across the entire coding region (Table 1, Figure 5). Original naming of these mutations (in square brackets [ ]) refers specifically to the leucine deletion [Leu], or for the conservative [422] T to C mutation, the nucleotide position from the start of exon 2. GDF9B mutations which changed amino acids were named [S1] and [S2]. Table 1 shows the relationship between (a) the original numbering system, (b) the coding nucleotide position in the full length coding sequence numbered from the first atg, (c) the position of the coding amino acid residue involved, and (d) the position of the coding residue (if any) within the mature coding sequence, numbering from the first amino acid after the furin protease processing site. For example mutation [S2] is a G to T nucleotide substitution at coding nucleotide 1100 of GDF9B which corresponds to an serine (Ser) residue changing to an isoleucine (Ile) residue at coding residue 367 of the full length unprocessed protein, or residue 99 of the processed mature peptide.

The first of these four polymorphisms (Table 1) is a previously-reported leucine deletion polymorphism [leu] in the predicted signal sequence (Galloway *et al.*, 2000) whereby some sheep have two leucine codons (CTT) at this position and some sheep have only one. This polymorphism has been shown to be unrelated to fertility and ovulation rate in Inverdale sheep (Galloway *et al.*, 2000). One other nucleotide change, [422], does not result in an altered amino acid (nucleotide position 747). The remaining two nucleotide changes ([S1] and [S2])give rise to more critical changes in

the protein (Figure 2, Figure 5). The [S1] C to T change at nucleotide 718 introduces a premature stop codon (TAG) in the place of glutamic acid (Q, CAG) at amino acid residue 239 of the unprocessed protein, which presumably results in complete loss of GDF9B function. The [S2] G to T change at nucleotide 1100 changes the serine residue at amino acid 99 of the mature active protein (residue 367 of the unprocessed protein) to an isoleucine, thereby substituting an uncharged polar group with a nonpolar group. The nucleotide and amino acid changes are illustrated in Figure 2 and Figure 5.

Screening for mutations in more animals

Initial sequencing of a smaller number of animals from each family identified the [74], [80], [324], [714], and [787] nucleotide changes in GDF9, and the [S1], [S2] and [422] changes in GDF9B. Forced RFLP (restriction fragment length polymorphism) assays to detect the specific SNPs were developed for [324], [714], [787] (GDF9) and for [S1] and [S2] (GDF9B), and these assays were carried out on larger numbers of animals (Table 2). Subsequent sequencing of full length GDF9 and GDF9B in more animals revealed the [581] and [597] nucleotide changes in GDF9 in the Cambridge sheep but not the F700 Belclares. Independent SSCP analysis identified the [E1] polymorphism in exon 1 of GDF9 in one ram, and this was also screened through further animals. [E1] was found to be associated with the wildtype alleles in this ram and his backcross progeny, and not associated with ovulation rate .

Homozygous mutations relate to sterility

Presence or absence of each of these nucleotide changes was analysed in relation to sterility or fertility in all of the animals tested, revealing that only the [787] change in GDF9 and the [S1] and [S2] changes in GDF9B contributed to infertility. Female sheep which are homozygous for [787] are sterile; female sheep which are homozygous for [S1], or homozygous for [S2] are sterile; female sheep which are

heterozygous for [S1] and [S2] simultaneously (whereby each each chromosome carries a different GDF9B mutation) are sterile. Figures 3a and 3b show small pedigrees illustrating what is seen in the larger set of animals.

Figure 3a illustrates a F700 Belclare pedigree. The sire R830 carries the GDF9B [S2] mutation on his X chromosome and the GDF9 [787] mutation on chromosome 5, but does not have the GDF9B [S1] mutation. Dam 9704 carries a single copy of the GDF9B [S1] mutation on her X chromosome and their two female offspring (930458 and 930459) are sterile since they have inherited inactive copies of GDF9B from both parents. Dam 8783 carries a single copy of the GDF9 [787] mutation on chromosome 5 and the female offspring of her mating with sire R830 are infertile and are homozygous for the GDF9 [787] mutation. Their infertility cannot be explained by GDF9B mutations. Offspring 930810 and 948302 are not homozygous for any of these mutations and hence are fertile. All three functional mutations ([S1], [S2], and [787]) were seen in the F700 Belclare flock (Table 2).

Figure 3b illustrates two Cambridge pedigrees. The sire 962101 carries the GDF9B [S1] mutation on his X chromosome and the GDF9 [787] mutation on chromosome 5, but does not have the GDF9B [S2] mutation. Dam 962152 carries a single copy of the [S1] mutation on her X chromosome and a single copy of the [787] mutation on chromosome 5. Their two female offspring (997634 and 997635) are sterile and have inherited inactive copies of both GDF9B ([S1]) and GDF9 ([787]) from both parents. Dam 976234 only carries a single copy of the [S1] mutation and one female offspring (997553) is infertile, having inherited inactive copies of GDF9B ([S1]) form both parents, whereas 997552 is fertile. Sire 930142 is homozygous for the GDf9 [787] mutation and carries the GDF9B [S1] mutation on his X chromosome, whereas dam 8874 is only heterozygous for the GDF9 [787] mutation and carries no GDF9B mutation. Their daughter (948093) has inherited two copies of the GDF9 [787] mutation and is sterile even though she also heterozygous for the GDF9B [S1]

mutation which she inherited from her sire. The [S2] mutation was not seen in any animals tested from the Cambridge flock (Table 2).

Among the animals tested for these changes we found fertile animals homozygous for GDF9 [324] and [714] and conclude that neither of those changes result in disruption of the genes sufficient to cause sterility. We also found animals which were heterozygous for GDF9 and GDF9B mutations together, and these animals were not sterile.

Structural effects of mutations on activity

Sufficient structural data has been obtained for members of the TGFβ superfamily to provide information about the likely effects of each of the three mutations ([S1], [S2] and [787]) on the biological activity of GDF9 and GDF9B, and hence explain the association with sterility. Structures have been reported for TGF β1 (Hinck *et al.*, 1996), TGF-β2 (Daopin *et al.*, 1992; Schlunegger and Grutter, 1992; Schlunegger and Grutter, 1993), TGF-β3 (Mittl *et al.*, 1996), BMP7/OP1 (Griffith *et al.*, 1996) and BMP2 (Scheufler *et al.*, 1999). Receptor binding structures have also been reported for BMP2 with the BRIA receptor binding ectodomain (Kirsch *et al.*, 2000a) and for TGF-β3 with ecTβR2 receptor binding ectodomain (Hart *et al.*, 2002).

The [S1] mutation results in premature termination of GDF9B protein prior to the mature active protein processing site. This mutation would result in no mature protein being translated, and is an even more severe effect than the Hanna mutation (Galloway et al., 2000) which results in infertility in sheep. The GDF9B [S2] mutation changes an uncharged polar serine residue (residue 99 of mature GDF9B) which is conserved across most members of the TGFβ superfamily, to a non-polar isoleucine (Figure 6). This serine (and the nearby conserved leucine) has been shown to be essential for receptor binding by structural and site-directed mutagenesis studies of BMP2 (Kirsch *et al.*, 2000b). In F700 Belclare sheep it appears that this mutation abolishes biological

activity of GDF9B, presumably by affecting receptor binding. The GDF9 [787] mutation changes an uncharged polar serine residue (residue 77 of mature GDF9) to a non-polar phenylalanine in a region of the molecule which is involved in dimerisation. This change occurs only three residues away from a conserved histidine (H80) of the mature GDF9 peptide (Figure 6). In BMP7 this conserved histidine exhibits hydrogen bonding to three residues of the paired molecule in the BMP7 dimer (Griffith *et al.*, 1996) and TGFβ3 (Mittl *et al.*, 1996). GDF9 lacks the interchain disulphide bond which forms a covalent link between both monomers of the biologically active dimer that is found in most other members of the TGFβ superfamily. This makes it likely that in GDF9 the hydrogen bonds between monomers are even more critical for maintaining dimer stability, and suggests that the GDF9 [787] mutation could be abolishing biological activity by disrupting dimerisation.

Heterozygous animals have increased ovulation rate

Irish Cambridge and F700 Belclare sheep have increased ovulation rates as well as infertility (Hanrahan, 1996). We examined ovulation rate data that was available for the fertile ewes which had been genotyped for the [S1], [S2] and [787] mutations (Table 3). Heterozygous carriers of mutations in GDF9B (either [S1] or [S2]) show increased ovulation rates similar to those seen in Inverdale and Hanna sheep (Davis *et al.*, 2001). Interestingly we see here for the first time that sheep heterozygous for the GDF9 [787] mutation also show increased ovulation rates and this increase is larger than for heterozygous carriers of GDF9B mutations (Table 3). It also appears that animals which are heterozygous for both a GDF9B mutation and a GDF9 mutation have an even higher ovulation rate, and this effect is probably additive. The effects of GDF9 [787] and GDF9B [S1] are additive in the Cambridge where the number of ewes in each genotype class was reasonably balanced and the effect of one copy of GDF9 [787] was twice as large as one copy of GDF9B [S1]. In the F700 Belclare the effect of one copy of GDF9B [S2] is close to the effect for [S1] in the Cambridge. The

evidence indicates that the effect of GDF9 [787] in the F700 Belclare is if similar magnitude as in the Cambridge and that the combination of GDF9B [S2] and GDF9 [787] give an ovulation rate that is similar to that from GDF9B [S1] and GDF9 [787] in Cambridge. Progeny test data for a set of nine F700 Belclare rams with genotype information for GDF9 and GDF9B mutations provides evidence that the effect of [S1] is 0.6 ; [S2] is 1.1 and GDF9 [787] is 1.8 . This suggests that the effect of [S2] on ovulation rate is greater than the effect of [S1].

Effect of immunising sheep against GDF9

In addition to the genotype effects above we have also shown that both long-term active immunisation and short-term passive immunisation of sheep with GDF9 causes sterility and/or abnormal corpus luteum function, and this finding provides additional evidence that a homozygous GDF9 mutant phenotype is one of sterility.

Repeated laparoscopy of ewes which had been in estrus and which were then were actively immunised against KLH (control, n = 9) or KLH conjugated to GDF9 peptide (n = 10) showed that no ewe actively immunised with GDF9 peptide showed cyclic estrous behaviour (Table 4). High (normal) progesterone concentrations was only seen in one or two samples (Figure 7), and most times when corpora lutea and/or luteal-like structures are observed following long-term immunisation against KLH conjugated to GDF9 peptide, progesterone concentrations are abnormal. In addition many of the ewes did not have any visible antral follicles at laparoscopy or at ovarian collection. Figure 7 also shows data for ewes immunised against GDF9B/BMP15 which we have shown previously (refer WO 01/96393), and which is included here for comparison with GDF9.

Passive immunisation using 100 ml of GDF9 peptide antiplasma can cause abnormal luteal function within 30 days of administration of the antiplasma (Figure 8). There were no differences in ovulation rates among the groups before administration of the

antiplasma. Administration of antiplasma against GDF9 peptide 4 days before induction of the follicular phase did not affect ovulation rate. However, at laparoscopy the corpora lutea of two of the animals treated with GDF9 antiplasma appeared smaller than normal. In addition, the overall mean concentration of progesterone during the subsequent luteal phase was less ($P<0.05$) than that observed in the control animals (Figure 8). This was the result of the progesterone concentrations being normal in two of the animals but in the other three animals, the post ovulation rise in progesterone was delayed even though luteolysis occurred at the normal time. Figure 8 also shows data for passive immunisation with GDF9B/BMP15 which we have shown previously (refer WO 01/96393), and which is included here for comparison with GDF9.

In another new experiment we show that short-term active immunisation of sheep with GDF9 or GDF9B can mimic the heterozygous effects of mutations in these genes. Short-term immunisation using milder adjuvant than in the previous experiments (2 immunisations in DEAE-Dextran adjuvant), with either KLH conjugated to GDF9 peptide or with KLH conjugated to GDF9B peptide, acts to increase ovulation rate in the animals which ovulated (Table 5). Moreover, more corpora lutea were evident in animals treated with GDF9 peptide than those treated with GDF9B.

CONCLUSIONS

These findings provide the first evidence that mutations in GDF9 and GDF9B are associated with the reproductive effects seen in the Cambridge and Belclare breeds of sheep. The increased ovulation rate and sterility phenotypes in these animals can be explained by the presence of heterozygous mutations and homozygous mutations, respectively, in these genes.

The two new mutations in GDF9B described here support the evidence from previous descriptions of mutations in this gene in sheep (Galloway *et al.*, 2000). Inactivating mutations in GDF9B cause increased ovulation rate and infertility in a dosage

dependent manner. The serine to isoleucine change in carriers of the [S2] mutation supports the notion that small perturbations of protein structure within the GDF9B mature peptide have serious consequences in protein activity.

The discovery of an inactivating mutation in GDF9 associated with infertility and increased ovulation rate in sheep is the first evidence that GDF9 is also important for increasing ovulation rate. Although a knockout mutation of GDF9 in mice has been shown to cause infertility, no effects for GDF9 on increasing ovulation rate have been described. Our discovery shows that small pertubations of protein structure within the GDF9 mature peptide also have severe consequences on protein activity.

These new mutations in GDF9B and GDF9 together provide strong support for the likelihood that other amino acid changes in the receptor-binding and dimerisation domains, or regions of the protein that disrupt protein folding of the mature peptide are likely to have similar effects. In addition, the effects of a GDF9 mutation and a GDF9B mutation together in one animal appear to be additive, implying that GDF9 and GDF9B are working independently, and that combinations of both proteins can be used to alter ovarian function more effectively than by altering either GDF9 or GDF9B alone.

Our finding of abnormal luteal function following GDF9 immunisation has not been previously reported. Given that corpus luteum function is often abnormal in the GDF9 immunised animals (both passive and active immunisations) it is likely that the administration of supplementary GDF9 or GDF9B, or GDF9 or GDF9B antagonists may modify corpus luteum function. More importantly, the observation that an increased ovulation rate effect (ie as seen in heterozygous carriers of these inactivating mutations) can also be induced in sheep by short-term active immunisation against peptides of GDF9 and GDF9B provides new methods for altering ovarian function.

TABLE 1

Sequence variations in GDF9 and GDF9B within the Irish Cambridge and F700 Belclare F700 flocks

Gene	[original name]	nucl. change	coding nucl.(bp)	coding residue	mature residue	result
GDF9	[E1]	G-A	260	87		Arg (R) – His (H)
	[74]	C-T	471	157		unchanged Val (V)
	[80]	G-A	477	159		unchanged Leu (L)
	[324]	G-A	721	241		Glu (E) – Lys (K)
	[581]	A-G	978	326	8	unchanged Glu (E)
	[597]	G-A	994	332	14	Val (V) – Ile (I)
	[714]	G-A	1111	371	53	Val (V) – Met (M)
	[787]	**C-T**	**1184**	**395**	**77**	**Ser (S) – Phe (F)**
GDF9B	[Leu]	CTT del	28-30	10		Leu deletion
	[S1]	C-T	718	239		Gln (Q) – STOP
	[422]	T-C	747	249		unchanged Pro (P)
	[S2]	**G-T**	**1100**	**367**	**99**	**Ser (S) – Ile (I)**

Columns indicate the relationship between (a) the original naming system used for each polymorphism, (b) the nucleotide change, (c) the coding nucleotide position (in base pairs (bp)) in the full length coding sequence numbered from the first atg start codon, (d) the

position of the coding amino acid residue involved (starting from the first Met residue), and (e) the position of the coding residue (if any) within the mature coding sequence, numbering from the first amino acid after the furin protease processing site. For example mutation [581] is a G to A nucleotide substitution at coding nucleotide 978 of GDF9 which corresponds to an unchanged glutamate (Glu) coding residue 326 of the full length unprocessed protein, or residue 8 of the processed mature peptide. Polymorphisms associated with infertility and ovulation rate traits are in bold.

TABLE 2. Genotype analysis of nucleotide changes in GDF9B and GDF9 genes from Irish Cambridge and F700 Belclare sheep.

Numbers shown are the number of individuals carrying at least one copy of the given mutation with the total number of individuals genotyped indicated in brackets underneath. Genotypes were determined by specific SNP assay and/or sequencing.

* The [E1] polymorphism in exon 1 of GDF9 (see Table 1) was identified by SSCP analysis and was not tested in the same set of animals used for the above table.

	GDF 9B			GDF 9*						
	[S1]	[S2]	[422]	[74]	[80]	[324]	[581]	[597]	[714]	[787]
F700 Belclares	9 (83)	71 (86)	2 (13)	6 (10)	6 (10)	13 (29)	0 (10)	0 (10)	2 (19)	11 (86)
Cambridge	74 (129)	0 (131)	0 (9)	0 (9)	7 (9)	1 (26)	3 (9)	2 (9)	7 (24)	95 (126)

TABLE 3. Least squares means for ovulation rate of sheep carrying the different genotypes for GDF9 and GDF9B mutations

Genotype			*Breed*	
S1	**S2**	**GDF9**	**F700 Belclare**	**Cambridge**
0	0	0	1.92±0.277 (n = 11)	2.27±0.488 (n = 10)
0	0	1	2.67±0.895 (n = 1)	4.39±0.308 (n = 28)
0	1	0	3.26±0.184 (n = 32)	-
0	1	1	6.09±0.549 (n = 3)	-
1	0	0	2.69±0.475 (n = 4)	3.11±0.438 (n = 15)
1	0	1	-	5.77±0.270(n = 38)
Effect of GDF9B [S1]			0.77±0.537 (P = 0.16)	1.18±0.387 (P<0.01)
Effect of GDF9B [S2]			2.38±0.548 (P<0.01)	-
Effect of GDF9 [787]			1.79±0.548 (P<0.01)	2.35±0.392 (P<0.01)
Interaction			-2.08±1.096 (P = 0.06)	-0.55±0.774 (P = 0.41)

( ) = no. of ewes

TABLE 4. Proportions of ewes immunized against KLH, GDF9 peptide or GDF9B peptide in estrus at the time of first laparoscopy (1^{st}), with visible luteal structures at laparoscopy 3-4 weeks before collection (2^{nd}) and at ovarian collection (3rd).

Immunized Group	1^{st}	2^{nd}	3^{rd}
KLH	9/9	9/9	9/9
GDF9 peptide	2/10*	2/10*	3/10*
GDF9B peptide	1/10*	1/10*	1/10*

*Signifies a value that is different from the control (KLH) value (P<0.05)

TABLE 5. Short-term immunisation of sheep with GDF9 or GDF9B

Treatment (n = no. of animals)	Ovulation rate (mean ± sem)
KLH (14)	2.1 ± 0.1
GDF9 peptide (7)	3.1 ± 0.4**
GDF9B peptide (6)	2.8 ± 0.3*

Results are expressed as average ovulation rates ± standard error of the mean (sem). In the ovulation rate column asterisks indicate a value that is significantly different than the control (KLH) group (* < 0.05, ** < 0.01)

Aspects of the present invention have been described by way of example only and it should be appreciated that modifications and additions may be made thereto without departing from the scope thereof.

AGRESEARCH LIMITED;
TEAGASC RESEARCH CENTRE;
NATIONAL DE LA RECHERCHE
AGRONOMIQUE and NATIONAL
UNIVERSITY OF IRELAND
by their Attorneys

JAMES & WELLS

REFERENCES

Aaltonen J, Laitinen M, Vuojolainen K, Jaatinen R, Horelli-Kuitunen N, Seppa L, Louhio H, Tuuri T, Sjoberg J, Butzow R, Hovatta O, Dale L, Ritvos O: Human growth differentiation factor- 9 (GDF-9) and its novel homolog GDF-9B are expressed during early folliculogenesis. J Clin Endocrinol Metab 84: 2744-2750, 1999

Bodensteiner, K.J., Clay, C.M., Moeller, C.L. and Sawyer, H.R. (1999) Molecular cloning of the ovine growth/differentiation factor-9 gene and expression of growth/differentiation factor-9 in ovine and bovine ovaries. Biology of Reproduction 60, 381-386

Bodensteiner, K.J., McNatty, K.P., Clay, C.M., Moeller, C.L. and Sawyer, H.R. (2000) Expression of growth and differentiation factor-9 in the ovaries of fetal sheep homozygous or heterozygous for the Inverdale prolificacy gene (*FecX*^{I}). Biology of Reproduction 62: 1479-1485.

Choreo and Goodman, Acc. Chem. REs., 1993 26 266-273

Daopin S, Piez KA, Ogawa Y, Davies DR: Crystal structure of transforming growth factor-beta 2: an unusual fold for the superfamily. Science 257:369-373, 1992.

Davis, G.H., Bruce, G.D. & Dodds, K.G. (2001) Ovulation rate and litter size of prolific Inverdale (FecX^{I}) and Hanna (FecX^{H}) sheep. Proceedings of Association for the Advancement of Animal Breeding and Genetics 14: 175-178

Dong J, Albertini DF, Nishimori K, Kumar TR, Lu N, Matzuk MM: Growth differentiation factor-9 is required during early ovarian folliculogenesis. Nature 383:531-535, 1996

Dube JL, Wang P, Elvin J, Lyons KM, Celeste AJ, Matzuk MM: The bone morphogenetic protein 15 gene is X-linked and expressed in oocytes. Mol Endocrinol 12:1809-1817, 1998

Eckery, D.C., Whale, L.J., Lawrence, S.B., Wilde, K.A., McNatty, K.P. & Juengel, J.L. (2002). Expression of mRNA encoding growth differentiation factor 9 and bone morphogenetic protein 15 during follicular formation and growth in a marsupial, the brushtail possum (*Trichasurus vulpecula*). Molecular & Cellular Endocrinology. In press.

Freidinger, R.M., Perlow, D.S., Veber, D.F.,*J. Org. Chem. 1982, 59*, 104-109

Gallop, M.A., Barrett, R.W., Dower, W.J., Fodor, S.P.A. and Hogan, Jr., J.C. Nature Biotechnology,1997 15 328-330.

Galloway SM, McNatty KP, Cambridge LM, Laitinen MPE, Juengel JL, Jokiranta TS, McLaren RJ, Luiro K, Dodds KG, Montgomery GW, Beattie AE, Davis GH, Ritvos O: Mutations in an oocyte-derived growth factor gene (*BMP15*) cause increased ovulation rate and infertility in a dosage-sensitive manner. Nature Genetics 25:279-283, 2000.

Grapes, L. and Rothschild, M.F. (2002). BMP15 maps to the X chromosome in swine. Animal Genetics 33, 165-166.

Griffith DL, Keck PC, Sampath TK, Rueger DC, Carlson WD: Three-dimensional structure of recombinant human osteogenic protein 1: structural paradigm for the transforming growth factor beta superfamily. Proc Natl Acad Sci U S A 93:878-883, 1996.

Hart, P.J., Deep, S., Taylor, A.B., Shu, Z., Hinck, C.S. and Hinck, A.P. (2002). Crystal structure of the human TβR2 ectodomain-TGF-β3 complex. Nature Structural Biology 9: 203-208.

Hanrahan, J.P. (1984) Results on selection for increased litter size and ovulation rate in sheep. In Proc 2nd World Congress Sheep & Beef Cattle Breeding pp 483-493.

Hanrahan, J.P. (1989) Reproductive performance of the Belclare sheeep: a new composite breed. 40th Annual Meeting of EAAP Dublin, paper 6.5 4pp

Hanrahan, J.P. (1991) Evidence for single gene effects on ovulation rate in the Cambridge and Belclare breeds. In *Major Genes for Reproduction in Sheep* pp 93-102 Eds J.M. Elsen, L.Bodin and J. Thimonier, INRA, Paris.

Hanrahan, JP. (1996) Irish Journal of Agricultural and Food Research 35: 87-88

Hanrahan, J.P. and Owen, J.B (1985) Variation and repeatability of ovulation rate in Cambridge ewes. British Society of Animal Production 40: 529

Hinck, A.P., Archer, S.J., Qian, S.W., Roberts, A.B., Sporn, M.B., Weatherbee, J.A., Tsang, M.L., Lucas, R., Zhang, B-L., Wenker, J. and Torchia, D.A. (1996) Transforming growth factor β1: Three-dimensional structure in solution and comparison with the X-ray structure of transforming growth factor β2. Biochemistry 35: 8517-8534.

Jaatinen, R., Laitinen, M.P., Vuojolainen, K., Aaltonen, J., Louhio, H., Heikinheimo, K., Lehtonen, E. and Ritvos, O. (1999) Localisation of growth differentiation factor-9 (Gdf-9) mRNA and protein in rat ovaries and cDNA cloning of rat GDF-9 and its novel homolog GDF-9B.Mol Cell Endocrinol 156: 189-193

Kirsch, T., Sebald, W. and Dreyer, M.K. (2000a). Crystal structure of the BMP2-BRIA ectodomain complex. Nature Structural Biology 7: 492-496.

Kirsch, T., Nickel, J. and Sebald, W. (2000b). BMP-2 antagonists emerge from alterations in the low-affinity binding epitope for receptor BMPR-II. EMBO Journal 19; 3314-3324.

Laitinen M, Vuojolainen K, Jaatinen R, Ketola I, Aaltonen J, Lehtonen E, Heikinheimo M, Ritvos O: A novel growth differentiation factor-9 (GDF-9) related factor is co-expressed with GDF-9 in mouse oocytes during folliculogenesis. Mech Dev 78:135-140, 1998

Maddox, J.F., Davies, K.P., Crawford, A.M., , Vaiman, D., Cribiu, E.P., Freking, B.A., Beh, K.J., Cockett, N.E., Kang, N., Riffkin, C.D., Drinkwater, R., Moore, S.S., Dodds, K.G., Lumsden, J.M., van Stijn, T.C., Phua, S.H., Adelson, D.L., Burkin, H.R., Broom, J.E., Buitkamp, J., Cambridge, L., Cushwa, W.T., Gerard, E., Galloway, S.M., Harrison, B., Hawken, R.J., Hiendleder, S., Henry, H.M., Medrano, J.F., Paterson, K.A., Schibler, L., Stone, R.T. and van Hest, B. (2001). An enhanced linkage map of the sheep genome comprising more than 1000 loci. Genome Research 11: 1275-1289.

McGrath SA, Esquela AF, Lee SJ : Oocyte-specific expression of growth differentiation factor-9. Mol Endocrinol 9:131-136, 1995.

McNatty K, Gibb M, Dobson C, Thurley, D, Findlay JK (1981). Australian Journal of Biological Science 34, 67-80

McPherron AC, Lee SJ: GDF-3 and GDF-9: two new members of the transforming growth factor-beta superfamily containing a novel pattern of cysteines. J Biol Chem 268:3444-3449, 1993

Mittl, P.R.E., Priestle, J.P., Cox, D.A., McMaster, G., Cerletti, N. and Grutter, M.G. (1996) The crystal structure of TGF-β3 and comparison to TGF-β3: Implications for receptor binding. Protein Science 5: 1261-1271.

Montgomery, G.W. and Sise J.A. (1990) Extraction of DNA form sheep white blood cells. NZ Journal of Agricultural Research 33, 437-441.

Mouse Genome Informatics database (20 May 2002)
http://www.informatics.jax.org/searches/homology_report.cgi

Nagai, U., Sato, K.*Tetrahedron Lett. 1985, 26,* 647-650.

Olson et al., J. Med. Chem., 1993 36 3039-3049.

Owen, J.B. 1991 *In* Prolific Sheep: Ed M.H. Fahmy.CAB International, Wallingford pp161 – 173

Sadighi, M., Bodensteiner, K.J., Beattie, A.E. and Galloway, S.M (2002). Genetic mapping of ovine growth differentiation factor 9 (*GDF9*) to sheep chromosome 5. Animal Genetics (*in press*).

Scheufler C, Sebald W, Hulsmeyer M: Crystal structure of human bone morphogenetic protein-2 at 2.7 A resolution. J Mol Biol 287:103-115, 1999

Schlunegger MP, Grutter MG: An unusual feature revealed by the crystal structure at 2.2 A resolution of human transforming growth factor-beta 2. Nature 358:430-434, 1992.

Schlunegger MP, Grutter MG: Refined crystal structure of human transforming growth factor beta 2 at 1.95 A resolution. J Mol Biol 231:445-458, 1993.

Smythe, M.L., von Itzstein, M., *J. Am. Chem. Soc. 1994, 116,* 2725-2733

Yan, C., Wang, P., DeMayo, J., DeMayo, F.J., Elvin, J., Carino, C., Prasad, S.V., Skinner, S.S., Dunbar, B.S., Dube, J.L., Celeste, A.J. and Matzuk, M.M. (2001) Synergistic roles of bone morphogenetic protein 15 and growth differentiation factor 9 in ovarian function. Molecular Endocrinology 15: 854-866.

SEQUENCE LISTING

<110> Galloway, Susan May
Davis, George Henry
Gregan, Scott Michael
Hanrahan, James Patrick
Juengal, Jennifer Lee
McNatty, Kenneth Pattrick
Mulsant, Philippe
Powell, Richard Patrick

<120> New sequences for altering mammalian ovarian function and ovulation rate

<130>

<160>

<170> PatentIn version 3.0

<210> 1
<211> 1879
<212> DNA
<213> Ovis aries

<220>
<221> 5'UTR
<222> (1)..(121)

<220>
<221> misc_feature
<222> (122)..(124)
<223> atg start codon.

<220>

<221> CDS
<222> (122)..(518)

<220>
<221> CDS
<222> (839)..(1800)

<220>
<221> Intron
<222> (519)..(838)
<223> n at 709 represents remainder of approx 900 bp unsequenced of the approx 1.1 kb intron

<220>
<221> mat_peptide
<222> (1396)..()

<220>
<221> misc_feature
<222> (709)..()
<223> n represents approximately 900 bp of unsequenced intron.

<220>
<221> misc_feature
<222> (1801)..(1803)
<223> taa stop codon.

<220>
<221> 3'UTR
<222> (1804)..(1879)

<220>
<221> mutation
<222> (1624)..(1626)

23> c to t at 1625 in [787] sheep changing tct serine codon to ttt phenylalanine

<400> 1

```
gaattgaacc tagcccaccc acacacctaa agtttattta agagaccaac cgaggctctt      60

cctggttttt aggaagaaga ctggtatggg gaaatgtgtt ccttgctaat tcttccaagc     120

c atg gcg ctt ccc  aac aaa ttc ttc ctt  tgg ttt tgc tgc ttt  gcc      166
  Met Ala Leu Pro  Asn Lys Phe Phe Leu  Trp Phe Cys Cys Phe  Ala
              -315                 -310                 -305

tgg ctc tgt ttt  cct att agc ctt gat  tct ctg cct tct agg  gga      211
Trp Leu Cys Phe  Pro Ile Ser Leu Asp  Ser Leu Pro Ser Arg  Gly
          -300                 -295                 -290

gaa gct cag att  gta gct agg act gcg  ttg gaa tct gag gct  gag      256
Glu Ala Gln Ile  Val Ala Arg Thr Ala  Leu Glu Ser Glu Ala  Glu
          -285                 -280                 -275

act tgg tcc ttg  ctg aac cat tta ggt  ggg aga cac aga cct  ggt      301
Thr Trp Ser Leu  Leu Asn His Leu Gly  Gly Arg His Arg Pro  Gly
          -270                 -265                 -260

ctc ctt tcc cct  ctc tta gag gtt ctg  tat gat ggg cac ggg  gaa      346
Leu Leu Ser Pro  Leu Leu Glu Val Leu  Tyr Asp Gly His Gly  Glu
          -255                 -250                 -245

ccc ccc agg ctg  cag cca gat gac aga  gct ttg cgc tac atg  aag      391
Pro Pro Arg Leu  Gln Pro Asp Asp Arg  Ala Leu Arg Tyr Met  Lys
          -240                 -235                 -230

agg ctc tat aag  gca tac gct acc aag  gag ggg acc cct aaa  tcc      436
Arg Leu Tyr Lys  Ala Tyr Ala Thr Lys  Glu Gly Thr Pro Lys  Ser
          -225                 -220                 -215

aac aga cgc cac  ctc tac aac act gtt  cgg ctc ttc acc ccc  tgt      481
Asn Arg Arg His  Leu Tyr Asn Thr Val  Arg Leu Phe Thr Pro  Cys
          -210                 -205                 -200

gct cag cac aag  cag gct cct ggg gac  ctg gcg gca g gtgtgtagga      528
Ala Gln His Lys  Gln Ala Pro Gly Asp  Leu Ala Ala
          -195                 -190

gcagattggt taatgggtgg agggaagaag aaagaccttt ttgcatttca gttacataaa    588

ggagttggcc ctgctccttg acttgcattt tactttgcat ggtactcaat atccaaacaa    648

acctggtgct tgatcttact tactgtttat tcctaatggc ctcatgggtt gatgtaggct    708

natcccaccc tgacgtttaa ggcttgagaa tgtggggaga aaagggacag aagcacattc    768

tgaggtactg attccttgat ttgacttcct gttacatatg gcattactgt tggattgttt    828

ttcttctcag ga  acc  ttt cca tca gtg gat  ctg ctg ttt aac ctg  gat    876
          Gly Thr  Phe Pro Ser Val Asp  Leu Leu Phe Asn Leu  Asp
              -185                 -180                 -175

cgt gtt act gtt  gtg gaa cat tta ttc  aag tca gtc ttg ctg  tat      921
Arg Val Thr Val  Val Glu His Leu Phe  Lys Ser Val Leu Leu  Tyr
          -170                 -165                 -160

act ttc aac aac  tcc att tct ttt ccc  ttt cct gtt aaa tgt  ata      966
Thr Phe Asn Asn  Ser Ile Ser Phe Pro  Phe Pro Val Lys Cys  Ile
          -155                 -150                 -145
```

```
aac ctg gtg  ata aaa gag cca gag  ttt tct agc aag act  ctc        1011
Asn Leu Val  Ile Lys Glu Pro Glu  Phe Ser Ser Lys Thr  Leu
            -140                 -135                 -130

cct aga gct cca  tac tca ttt acc tat  aac tca cag ttt gaa  ttt    1056
Pro Arg Ala Pro  Tyr Ser Phe Thr Tyr  Asn Ser Gln Phe Glu  Phe
            -125                 -120                 -115

aga aag aaa tac  aaa tgg atg gag att  gat gtg acg gct cct  ctt    1101
Arg Lys Lys Tyr  Lys Trp Met Glu Ile  Asp Val Thr Ala Pro  Leu
            -110                 -105                 -100

gag cct ctg gtg gcc tcc cac aag agg aat att cac atg tct gta aat  1149
Glu Pro Leu Val Ala Ser His Lys Arg Asn Ile His Met Ser Val Asn
            -95                  -90                  -85

ttt aca tgt gcg gaa gac cag ctg cag cat cct tca gcg cgg gac agc  1197
Phe Thr Cys Ala Glu Asp Gln Leu Gln His Pro Ser Ala Arg Asp Ser
        -80                  -75                  -70

ctg ttt aac atg act ctt ctc gta gcg ccc tca ctg ctt ttg tat ctg  1245
Leu Phe Asn Met Thr Leu Leu Val Ala Pro Ser Leu Leu Leu Tyr Leu
    -65                  -60                  -55

aac gac aca agt gct cag gct ttt cac agg tgg cat tcc ctc cac cct  1293
Asn Asp Thr Ser Ala Gln Ala Phe His Arg Trp His Ser Leu His Pro
-50                  -45                  -40                  -35

aaa agg aag cct tca cag ggt cct gac cag aag aga ggg cta tct gcc  1341
Lys Arg Lys Pro Ser Gln Gly Pro Asp Gln Lys Arg Gly Leu Ser Ala
                -30                  -25                  -20

tac ccc gtg gga gaa gaa gct gct gag ggt gta aga tcg tcc cgt cac  1389
Tyr Pro Val Gly Glu Glu Ala Ala Glu Gly Val Arg Ser Ser Arg His
            -15                  -10                  -5

cgc aga gac cag gag agt gcc agc tct gaa ttg aag aag cct ctg gtt  1437
Arg Arg Asp Gln Glu Ser Ala Ser Ser Glu Leu Lys Lys Pro Leu Val
    -1  1                5                   10

cca gct tca gtc aat ctg agt gaa tac ttc aaa cag ttt ctt ttt ccc  1485
Pro Ala Ser Val Asn Leu Ser Glu Tyr Phe Lys Gln Phe Leu Phe Pro
15                  20                  25                  30

cag aat gaa tgt gag ctc cat gac ttt aga ctt agc ttt agt cag ctg  1533
Gln Asn Glu Cys Glu Leu His Asp Phe Arg Leu Ser Phe Ser Gln Leu
                35                  40                  45

aag tgg gac aac tgg att gtg gcc cca cac aaa tac aac cct cga tac  1581
Lys Trp Asp Asn Trp Ile Val Ala Pro His Lys Tyr Asn Pro Arg Tyr
            50                  55                  60

tgt aaa ggg gac tgt ccc agg gcg gtc gga cat cgg tat ggc ttt ccg  1629
Cys Lys Gly Asp Cys Pro Arg Ala Val Gly His Arg Tyr Gly Phe Pro
        65                  70                  75

gtt cac acc atg gtg cag aac atc atc cat gag aaa ctt gac tcc tca  1677
Val His Thr Met Val Gln Asn Ile Ile His Glu Lys Leu Asp Ser Ser
    80                  85                  90

gtg cca aga cca tcc tgt gta cct gcc aag tat agc cct ttg agt gtt  1725
Val Pro Arg Pro Ser Cys Val Pro Ala Lys Tyr Ser Pro Leu Ser Val
95                  100                 105                 110

ttg gcc atc gag cct gat ggc tca atc gct tat aaa gaa tat gaa gat  1773
Leu Ala Ile Glu Pro Asp Gly Ser Ile Ala Tyr Lys Glu Tyr Glu Asp
                115                 120                 125
```

ata gcc act aag tgt acc tgt cgt taacagactc ctgtcaagta 1820
Ile Ala Thr Lys Cys Thr Cys Arg
130 135

aaaccatgag tgtcctggcc agtgtaaatg ccgcgcccct gtctatgcct ttgggagga 1879

<210> 2

<211> 453

<212> PRT

<213> Ovis aries

<220>

<221> misc_feature

<222> (122)..(124)

<223> atg start codon.

<220>

<221> misc_feature

<222> (709)..()

<223> n represents approximately 900 bp of unsequenced intron.

<220>

<221> misc_feature

<222> (1801)..(1803)

<223> taa stop codon.

<400> 2

Met Ala Leu Pro Asn Lys Phe Phe Leu Trp Phe Cys Cys Phe Ala
-315 -310 -305

Trp Leu Cys Phe Pro Ile Ser Leu Asp Ser Leu Pro Ser Arg Gly
-300 -295 -290

Glu Ala Gln Ile Val Ala Arg Thr Ala Leu Glu Ser Glu Ala Glu
-285 -280 -275

Thr Trp Ser Leu Leu Asn His Leu Gly Gly Arg His Arg Pro Gly
-270 -265 -260

Leu Leu Ser Pro Leu Leu Glu Val Leu Tyr Asp Gly His Gly Glu
-255 -250 -245

Pro Pro Arg Leu Gln Pro Asp Asp Arg Ala Leu Arg Tyr Met Lys
-240 -235 -230

Arg Leu Tyr Lys Ala Tyr Ala Thr Lys Glu Gly Thr Pro Lys Ser

```
                -225                -220                -215

Asn Arg Arg His  Leu Tyr Asn Thr Val  Arg Leu Phe Thr Pro  Cys
            -210                -205                -200

Ala Gln His Lys  Gln Ala Pro Gly Asp  Leu Ala Ala Gly Thr  Phe
            -195                -190                -185

Pro Ser Val Asp  Leu Leu Phe Asn Leu  Asp Arg Val Thr Val  Val
            -180                -175                -170

Glu His Leu Phe  Lys Ser Val Leu Leu  Tyr Thr Phe Asn Asn  Ser
            -165                -160                -155

Ile Ser Phe Pro  Phe Pro Val Lys Cys  Ile Cys Asn Leu Val  Ile
            -150                -145                -140

Lys Glu Pro Glu  Phe Ser Ser Lys Thr  Leu Pro Arg Ala Pro  Tyr
            -135                -130                -125

Ser Phe Thr Tyr  Asn Ser Gln Phe Glu  Phe Arg Lys Lys Tyr  Lys
            -120                -115                -110

Trp Met Glu Ile  Asp Val Thr Ala Pro  Leu Glu Pro Leu Val Ala Ser
            -105                -100                -95

His Lys Arg Asn Ile His Met Ser Val Asn Phe Thr Cys Ala Glu Asp
        -90                 -85                 -80

Gln Leu Gln His Pro Ser Ala Arg Asp Ser Leu Phe Asn Met Thr Leu
    -75                 -70                 -65

Leu Val Ala Pro Ser Leu Leu Leu Tyr Leu Asn Asp Thr Ser Ala Gln
-60                  -55                 -50                 -45

Ala Phe His Arg Trp His Ser Leu His Pro Lys Arg Lys Pro Ser Gln
                -40                 -35                 -30

Gly Pro Asp Gln Lys Arg Gly Leu Ser Ala Tyr Pro Val Gly Glu Glu
            -25                 -20                 -15

Ala Ala Glu Gly Val Arg Ser Ser Arg His Arg Arg Asp Gln Glu Ser
        -10                 -5                  -1  1

Ala Ser Ser Glu Leu Lys Lys Pro Leu Val Pro Ala Ser Val Asn Leu
5                   10                  15                  20

Ser Glu Tyr Phe Lys Gln Phe Leu Phe Pro Gln Asn Glu Cys Glu Leu
                25                  30                  35

His Asp Phe Arg Leu Ser Phe Ser Gln Leu Lys Trp Asp Asn Trp Ile
```

40 45 50

Val Ala Pro His Lys Tyr Asn Pro Arg Tyr Cys Lys Gly Asp Cys Pro
55 60 65

Arg Ala Val Gly His Arg Tyr Gly Phe Pro Val His Thr Met Val Gln
70 75 80

Asn Ile Ile His Glu Lys Leu Asp Ser Ser Val Pro Arg Pro Ser Cys
85 90 95 100

Val Pro Ala Lys Tyr Ser Pro Leu Ser Val Leu Ala Ile Glu Pro Asp
105 110 115

Gly Ser Ile Ala Tyr Lys Glu Tyr Glu Asp Met Ile Ala Thr Lys Cys
120 125 130

Thr Cys Arg
135

<210> 3

<211> 1362

<212> DNA

<213> Ovis aries

<220>

<221> misc_feature

<222> (1)..(3)

<223> atg start codon.

<220>

<221> CDS

<222> (1)..(1359)

<220>

<221> mat_peptide

<222> (955)..()

<220>

<221> mutation

<222> (1183)..(1185)

3> c to t 1184 in [787] sheep changing serine tct codon to ttt pheny lalanine

<220>

<221> misc_feature

<222> (1360)..(1362)

<223> taa stop codon.

<400> 3

```
atg gcg ctt ccc  aac aaa ttc ttc ctt  tgg ttt tgc tgc ttt  gcc        45
Met Ala Leu Pro  Asn Lys Phe Phe Leu  Trp Phe Cys Cys Phe  Ala
          -315                -310                -305

tgg ctc tgt ttt  cct att agc ctt gat  tct ctg cct tct agg  gga        90
Trp Leu Cys Phe  Pro Ile Ser Leu Asp  Ser Leu Pro Ser Arg  Gly
          -300                -295                -290

gaa gct cag att  gta gct agg act gcg  ttg gaa tct gag gct  gag       135
Glu Ala Gln Ile  Val Ala Arg Thr Ala  Leu Glu Ser Glu Ala  Glu
          -285                -280                -275

act tgg tcc ttg  ctg aac cat tta ggt  ggg aga cac aga cct  ggt       180
Thr Trp Ser Leu  Leu Asn His Leu Gly  Gly Arg His Arg Pro  Gly
          -270                -265                -260

ctc ctt tcc cct  ctc tta gag gtt ctg  tat gat ggg cac ggg  gaa       225
Leu Leu Ser Pro  Leu Leu Glu Val Leu  Tyr Asp Gly His Gly  Glu
          -255                -250                -245

ccc ccc agg ctg  cag cca gat gac aga  gct ttg cgc tac atg  aag       270
Pro Pro Arg Leu  Gln Pro Asp Asp Arg  Ala Leu Arg Tyr Met  Lys
          -240                -235                -230

agg ctc tat aag  gca tac gct acc aag  gag ggg acc cct aaa  tcc       315
Arg Leu Tyr Lys  Ala Tyr Ala Thr Lys  Glu Gly Thr Pro Lys  Ser
          -225                -220                -215

aac aga cgc cac  ctc tac aac act gtt  cgg ctc ttc acc ccc  tgt       360
Asn Arg Arg His  Leu Tyr Asn Thr Val  Arg Leu Phe Thr Pro  Cys
          -210                -205                -200

gct cag cac aag  cag gct cct ggg gac  ctg gcg gca gga acc  ttt       405
Ala Gln His Lys  Gln Ala Pro Gly Asp  Leu Ala Ala Gly Thr  Phe
          -195                -190                -185

cca tca gtg gat  ctg ctg ttt aac ctg  gat cgt gtt act gtt  gtg       450
Pro Ser Val Asp  Leu Leu Phe Asn Leu  Asp Arg Val Thr Val  Val
          -180                -175                -170

gaa cat tta ttc  aag tca gtc ttg ctg  tat act ttc aac aac  tcc       495
Glu His Leu Phe  Lys Ser Val Leu Leu  Tyr Thr Phe Asn Asn  Ser
          -165                -160                -155

att tct ttt ccc  ttt cct gtt aaa tgt  ata tgc aac ctg gtg  ata       540
Ile Ser Phe Pro  Phe Pro Val Lys Cys  Ile Cys Asn Leu Val  Ile
          -150                -145                -140

aaa gag cca gag  ttt tct agc aag act  ctc cct aga gct cca  tac       585
Lys Glu Pro Glu  Phe Ser Ser Lys Thr  Leu Pro Arg Ala Pro  Tyr
          -135                -130                -125
```

```
ttt acc tat  aac tca cag ttt gaa  ttt aga aag aaa tac  aaa          630
Phe Thr Tyr  Asn Ser Gln Phe Glu  Phe Arg Lys Lys Tyr  Lys
            -120                 -115                 -110

tgg atg gag att  gat gtg acg gct cct  ctt gag cct ctg gtg gcc tcc    678
Trp Met Glu Ile  Asp Val Thr Ala Pro  Leu Glu Pro Leu Val Ala Ser
            -105                 -100                 -95

cac aag agg aat att cac atg tct gta aat ttt aca tgt gcg gaa gac     726
His Lys Arg Asn Ile His Met Ser Val Asn Phe Thr Cys Ala Glu Asp
        -90                  -85                  -80

cag ctg cag cat cct tca gcg cgg gac agc ctg ttt aac atg act ctt     774
Gln Leu Gln His Pro Ser Ala Arg Asp Ser Leu Phe Asn Met Thr Leu
    -75                  -70                  -65

ctc gta gcg ccc tca ctg ctt ttg tat ctg aac gac aca agt gct cag     822
Leu Val Ala Pro Ser Leu Leu Leu Tyr Leu Asn Asp Thr Ser Ala Gln
-60                  -55                  -50                  -45

gct ttt cac agg tgg cat tcc ctc cac cct aaa agg aag cct tca cag     870
Ala Phe His Arg Trp His Ser Leu His Pro Lys Arg Lys Pro Ser Gln
                -40                  -35                  -30

ggt cct gac cag aag aga ggg cta tct gcc tac ccc gtg gga gaa gaa     918
Gly Pro Asp Gln Lys Arg Gly Leu Ser Ala Tyr Pro Val Gly Glu Glu
            -25                  -20                  -15

gct gct gag ggt gta aga tcg tcc cgt cac cgc aga gac cag gag agt     966
Ala Ala Glu Gly Val Arg Ser Ser Arg His Arg Arg Asp Gln Glu Ser
        -10                  -5                   -1  1

gcc agc tct gaa ttg aag aag cct ctg gtt cca gct tca gtc aat ctg     1014
Ala Ser Ser Glu Leu Lys Lys Pro Leu Val Pro Ala Ser Val Asn Leu
5                   10                   15                  20

agt gaa tac ttc aaa cag ttt ctt ttt ccc cag aat gaa tgt gag ctc     1062
Ser Glu Tyr Phe Lys Gln Phe Leu Phe Pro Gln Asn Glu Cys Glu Leu
                25                   30                   35

cat gac ttt aga ctt agc ttt agt cag ctg aag tgg gac aac tgg att     1110
His Asp Phe Arg Leu Ser Phe Ser Gln Leu Lys Trp Asp Asn Trp Ile
            40                   45                   50

gtg gcc cca cac aaa tac aac cct cga tac tgt aaa ggg gac tgt ccc     1158
Val Ala Pro His Lys Tyr Asn Pro Arg Tyr Cys Lys Gly Asp Cys Pro
        55                   60                   65

agg gcg gtc gga cat cgg tat ggc ttt ccg gtt cac acc atg gtg cag     1206
Arg Ala Val Gly His Arg Tyr Gly Phe Pro Val His Thr Met Val Gln
    70                   75                   80

aac atc atc cat gag aaa ctt gac tcc tca gtg cca aga cca tcc tgt     1254
Asn Ile Ile His Glu Lys Leu Asp Ser Ser Val Pro Arg Pro Ser Cys
85                  90                   95                  100

gta cct gcc aag tat agc cct ttg agt gtt ttg gcc atc gag cct gat     1302
Val Pro Ala Lys Tyr Ser Pro Leu Ser Val Leu Ala Ile Glu Pro Asp
                105                  110                  115

ggc tca atc gct tat aaa gaa tat gaa gat atg ata gcc act aag tgt     1350
Gly Ser Ile Ala Tyr Lys Glu Tyr Glu Asp Met Ile Ala Thr Lys Cys
            120                  125                  130

acc tgt cgt taa                                                    1362
Thr Cys Arg
        135
```

<210> 4
<211> 453
<212> PRT
<213> Ovis aries

<220>
<221> misc_feature
<222> (1)..(3)
<223> atg start codon.

<220>
<221> misc_feature
<222> (1360)..(1362)
<223> taa stop codon.

<400> 4

```
Met Ala Leu Pro  Asn Lys Phe Phe Leu  Trp Phe Cys Cys Phe  Ala
            -315                 -310                 -305

Trp Leu Cys Phe  Pro Ile Ser Leu Asp  Ser Leu Pro Ser Arg  Gly
            -300                 -295                 -290

Glu Ala Gln Ile  Val Ala Arg Thr Ala  Leu Glu Ser Glu Ala  Glu
            -285                 -280                 -275

Thr Trp Ser Leu  Leu Asn His Leu Gly  Gly Arg His Arg Pro  Gly
            -270                 -265                 -260

Leu Leu Ser Pro  Leu Leu Glu Val Leu  Tyr Asp Gly His Gly  Glu
            -255                 -250                 -245

Pro Pro Arg Leu  Gln Pro Asp Asp Arg  Ala Leu Arg Tyr Met  Lys
            -240                 -235                 -230

Arg Leu Tyr Lys  Ala Tyr Ala Thr Lys  Glu Gly Thr Pro Lys  Ser
            -225                 -220                 -215

Asn Arg Arg His  Leu Tyr Asn Thr Val  Arg Leu Phe Thr Pro  Cys
            -210                 -205                 -200

Ala Gln His Lys  Gln Ala Pro Gly Asp  Leu Ala Ala Gly Thr  Phe
            -195                 -190                 -185

Pro Ser Val Asp  Leu Leu Phe Asn Leu  Asp Arg Val Thr Val  Val
            -180                 -175                 -170
```

```
His Leu Phe Lys Ser Val Leu Leu Tyr Thr Phe Asn Asn Ser
-165                -160                -155

Ile Ser Phe Pro Phe Pro Val Lys Cys Ile Cys Asn Leu Val Ile
-150                -145                -140

Lys Glu Pro Glu Phe Ser Ser Lys Thr Leu Pro Arg Ala Pro Tyr
-135                -130                -125

Ser Phe Thr Tyr Asn Ser Gln Phe Glu Phe Arg Lys Lys Tyr Lys
-120                -115                -110

Trp Met Glu Ile Asp Val Thr Ala Pro Leu Glu Pro Leu Val Ala Ser
-105                 -100                -95

His Lys Arg Asn Ile His Met Ser Val Asn Phe Thr Cys Ala Glu Asp
-90                  -85                 -80

Gln Leu Gln His Pro Ser Ala Arg Asp Ser Leu Phe Asn Met Thr Leu
-75                  -70                 -65

Leu Val Ala Pro Ser Leu Leu Leu Tyr Leu Asn Asp Thr Ser Ala Gln
-60                  -55                 -50                 -45

Ala Phe His Arg Trp His Ser Leu His Pro Lys Arg Lys Pro Ser Gln
                -40                 -35                 -30

Gly Pro Asp Gln Lys Arg Gly Leu Ser Ala Tyr Pro Val Gly Glu Glu
            -25                 -20                 -15

Ala Ala Glu Gly Val Arg Ser Ser Arg His Arg Arg Asp Gln Glu Ser
        -10                 -5                  -1  1

Ala Ser Ser Glu Leu Lys Lys Pro Leu Val Pro Ala Ser Val Asn Leu
5                   10                  15                  20

Ser Glu Tyr Phe Lys Gln Phe Leu Phe Pro Gln Asn Glu Cys Glu Leu
                25                  30                  35

His Asp Phe Arg Leu Ser Phe Ser Gln Leu Lys Trp Asp Asn Trp Ile
            40                  45                  50

Val Ala Pro His Lys Tyr Asn Pro Arg Tyr Cys Lys Gly Asp Cys Pro
        55                  60                  65

Arg Ala Val Gly His Arg Tyr Gly Phe Pro Val His Thr Met Val Gln
    70                  75                  80

Asn Ile Ile His Glu Lys Leu Asp Ser Ser Val Pro Arg Pro Ser Cys
85                  90                  95                  100
```

Pro Ala Lys Tyr Ser Pro Leu Ser Val Leu Ala Ile Glu Pro Asp
105 110 115

Gly Ser Ile Ala Tyr Lys Glu Tyr Glu Asp Met Ile Ala Thr Lys Cys
120 125 130

Thr Cys Arg
135

<210> 5

<211> 168

<212> DNA

<213> Ovis aries

<220>

<221> CDS

<222> (1)..(168)

<220>

<221> mutation

<222> (82)..(84)

<223> c to t at 83 in [787] sheep changing tct serine codon to ttt phenylalanine

<400> 5

```
aac tgg att gtg gcc cca cac aaa tac aac cct cga tac tgt aaa ggg       48
Asn Trp Ile Val Ala Pro His Lys Tyr Asn Pro Arg Tyr Cys Lys Gly
1               5                  10                  15

gac tgt ccc agg gcg gtc gga cat cgg tat ggc ttt ccg gtt cac acc       96
Asp Cys Pro Arg Ala Val Gly His Arg Tyr Gly Phe Pro Val His Thr
            20                  25                  30

atg gtg cag aac atc atc cat gag aaa ctt gac tcc tca gtg cca aga      144
Met Val Gln Asn Ile Ile His Glu Lys Leu Asp Ser Ser Val Pro Arg
        35                  40                  45

cca tcc tgt gta cct gcc aag tat                                     168
Pro Ser Cys Val Pro Ala Lys Tyr
    50                  55
```

<210> 6

<211> 56

<212> PRT

<213> Ovis aries

<400> 6

Asn Trp Ile Val Ala Pro His Lys Tyr Asn Pro Arg Tyr Cys Lys Gly
1 5 10 15

Asp Cys Pro Arg Ala Val Gly His Arg Tyr Gly Phe Pro Val His Thr
20 25 30

Met Val Gln Asn Ile Ile His Glu Lys Leu Asp Ser Ser Val Pro Arg
35 40 45

Pro Ser Cys Val Pro Ala Lys Tyr
50 55

<210> 7

<211> 1665

<212> DNA

<213> Ovis aries

<220>

<221> 5'UTR

<222> (1)..(252)

<220>

<221> misc_feature

<222> (253)..(255)

<223> atg start codon.

<220>

<221> CDS

<222> (253)..(577)

<220>

<221> CDS

<222> (774)..(1165)

<220>

<221> Intron

2> (578)..(773)
<223> n at 685 represents remainder of approx 5.2 kb intron.

<220>
<221> misc_feature
<222> (1253)..(1255)
<223> position of first codon of mature peptide in wildtype sheep.

<220>
<221> misc_feature
<222> (685)..()
<223> n represents approx 5.2 kb of intron.

<220>
<221> misc_feature
<222> (1628)..(1630)
<223> tga stop codon in wildtype sheep.

<220>
<221> 3'UTR
<222> (1631)..(1665)

<220>
<221> mutation
<222> (1166)..(1168)
<223> c to t at 1166 of [S1] sheep changes cag glutamine codon to tag STOP

<400> 7
catgctgcct tgtcccacct gctgtttctg tttgtttgat gcaaagagga caatttagaa 60

gacctctttt tggttcagga gatcctacca gaggaagaaa cataggacct gcctgccagc 120

ctttcatttt tccttgccct atcctttgtg gtagtggagc ctggatgctg ttacccatgt 180

aaaaggaaag gtttaaagcg ttatcctttg ggcttttatc agaacatgtt gctgaacacc 240

aagcttttca ag atg gtc ctc ctg agc atc ctt aga atc ctt ctt tgg gga 291
Met Val Leu Leu Ser Ile Leu Arg Ile Leu Leu Trp Gly
1 5 10

```
gtg ctt ttt atg gaa cat agg gtc caa atg aca cag gta ggg cag      339
Val Leu Phe Met Glu His Arg Val Gln Met Thr Gln Val Gly Gln
 15                  20                  25

ccc tct att gcc cac ctg cct gag gcc cct acc ttg ccc ctg att cag  387
Pro Ser Ile Ala His Leu Pro Glu Ala Pro Thr Leu Pro Leu Ile Gln
30                  35                  40                  45

gag ctg cta gaa gaa gcc cct ggc aag cag cag agg aag ccg cgg gtc  435
Glu Leu Leu Glu Glu Ala Pro Gly Lys Gln Gln Arg Lys Pro Arg Val
                50                  55                  60

tta ggg cat ccc tta cgg tat atg ctg gag ctg tac cag cgt tca gct  483
Leu Gly His Pro Leu Arg Tyr Met Leu Glu Leu Tyr Gln Arg Ser Ala
            65                  70                  75

gac gca agt gga cac cct agg gaa aac cgc acc att ggg gcc acc atg  531
Asp Ala Ser Gly His Pro Arg Glu Asn Arg Thr Ile Gly Ala Thr Met
        80                  85                  90

gtg agg ctg gtg agg ccg ctg gct agt gta gca agg cct ctc aga g    577
Val Arg Leu Val Arg Pro Leu Ala Ser Val Ala Arg Pro Leu Arg
    95                  100                 105

gtgagttatc atactatatt gttctggtgg gagggggga gaaaatgggg aagaaaagtg  637

tagaaaaaag tggatctgtc agttttctgt caggcttcac attgcctnca gtttgtactg  697

agcaggtctg ttagagagac taaggctagg atataagaag ctaacgcttt gctcttgttc  757

cctcttacta atgcag gc  tcc tgg cac ata cag acc ctg gac ttt cct ctg  808
                 Gly Ser Trp His Ile Gln Thr Leu Asp Phe Pro Leu
                 110                 115                 120

aga cca aac cgg gta gca tac caa cta gtc aga gcc act gtg gtt tac  856
Arg Pro Asn Arg Val Ala Tyr Gln Leu Val Arg Ala Thr Val Val Tyr
                125                 130                 135

cgc cat cag ctt cac cta act cat tcc cac ctc tcc tgc cat gtg gag  904
Arg His Gln Leu His Leu Thr His Ser His Leu Ser Cys His Val Glu
            140                 145                 150

ccc tgg gtc cag aaa agc cca acc aat cac ttt cct tct tca gga aga  952
Pro Trp Val Gln Lys Ser Pro Thr Asn His Phe Pro Ser Ser Gly Arg
        155                 160                 165

ggc tcc tca aag cct tcc ctg ttg ccc aaa act tgg aca gag atg gat  1000
Gly Ser Ser Lys Pro Ser Leu Leu Pro Lys Thr Trp Thr Glu Met Asp
    170                 175                 180

atc atg gaa cat gtt ggg caa aag ctc tgg aat cac aag ggg cgc agg  1048
Ile Met Glu His Val Gly Gln Lys Leu Trp Asn His Lys Gly Arg Arg
185                 190                 195                 200

gtt cta cga ctc cgc ttc gtg tgt cag cag cca aga ggt agt gag gtt  1096
Val Leu Arg Leu Arg Phe Val Cys Gln Gln Pro Arg Gly Ser Glu Val
                205                 210                 215

ctt gag ttc tgg tgg cat ggc act tca tca ttg gac act gtc ttc ttg  1144
Leu Glu Phe Trp Trp His Gly Thr Ser Ser Leu Asp Thr Val Phe Leu
            220                 225                 230

tta ctg tat ttc aat gac act tagagtgttc agaagaccaa acctctccct     1195
Leu Leu Tyr Phe Asn Asp Thr
        235

aaaggcctga aagagtttac agaaaaagac ccttctcttc tcttgaggag ggctcgtcaa  1255
```

ggcagta ttgcatcgga agttcctggc ccctccaggg agcatgatgg gcctgaaagt 1315
aaccagtgtt ccctccaccc ttttcaagtc agcttccagc agctgggctg ggatcactgg 1375
atcattgctc cccatctcta taccccaaac tactgtaagg gagtatgtcc tcgggtacta 1435
cactatggtc tcaattctcc caatcatgcc atcatccaga accttgtcag tgagctggtg 1495
gatcagaatg tccctcagcc ttcctgtgtc ccttataagt atgttcccat tagcatcctt 1555
ctgattgagg caaatgggag tatcttgtac aaggagtatg agggtatgat tgcccagtcc 1615
tgcacatgca ggtgacggca aaggtgcagc tagctcaggt ttcccaagaa 1665

<210> 8
<211> 239
<212> PRT
<213> Ovis aries

<220>
<221> misc_feature
<222> (253)..(255)
<223> atg start codon.
<220>
<221> misc_feature
<222> (1253)..(1255)
<223> position of first codon of mature peptide in wildtype sheep.
<220>
<221> misc_feature
<222> (685)..()
<223> n represents approx 5.2 kb of intron.
<220>
<221> misc_feature
<222> (1628)..(1630)
<223> tga stop codon in wildtype sheep.
<400> 8

Met Val Leu Leu Ser Ile Leu Arg Ile Leu Leu Trp Gly Leu Val Leu
1 5 10 15

Phe Met Glu His Arg Val Gln Met Thr Gln Val Gly Gln Pro Ser Ile
20 25 30

His Leu Pro Glu Ala Pro Thr Leu Pro Leu Ile Gln Glu Leu Leu
35 40 45

Glu Glu Ala Pro Gly Lys Gln Gln Arg Lys Pro Arg Val Leu Gly His
50 55 60

Pro Leu Arg Tyr Met Leu Glu Leu Tyr Gln Arg Ser Ala Asp Ala Ser
65 70 75 80

Gly His Pro Arg Glu Asn Arg Thr Ile Gly Ala Thr Met Val Arg Leu
85 90 95

Val Arg Pro Leu Ala Ser Val Ala Arg Pro Leu Arg Gly Ser Trp His
100 105 110

Ile Gln Thr Leu Asp Phe Pro Leu Arg Pro Asn Arg Val Ala Tyr Gln
115 120 125

Leu Val Arg Ala Thr Val Val Tyr Arg His Gln Leu His Leu Thr His
130 135 140

Ser His Leu Ser Cys His Val Glu Pro Trp Val Gln Lys Ser Pro Thr
145 150 155 160

Asn His Phe Pro Ser Ser Gly Arg Gly Ser Ser Lys Pro Ser Leu Leu
165 170 175

Pro Lys Thr Trp Thr Glu Met Asp Ile Met Glu His Val Gly Gln Lys
180 185 190

Leu Trp Asn His Lys Gly Arg Arg Val Leu Arg Leu Arg Phe Val Cys
195 200 205

Gln Gln Pro Arg Gly Ser Glu Val Leu Glu Phe Trp Trp His Gly Thr
210 215 220

Ser Ser Leu Asp Thr Val Phe Leu Leu Leu Tyr Phe Asn Asp Thr
225 230 235

<210> 9

<211> 1182

<212> DNA

<213> Ovis areis

<220>

<221> misc_feature

<222> (1)..(3)

<223> atg start codon.

<220>

<221> mutation

<222> (718)..(720)

<223> c to t at 718 of [S1] sheep changes cag glutamine codon to tag STOP.

<220>

<221> CDS

<222> (1)..(717)

<220>

<221> misc_feature

<222> (805)..(807)

<223> first codon of mature peptide in wildtype sheep.

<220>

<221> misc_feature

<222> (1180)..(1182)

<223> tga stop codon.

<400> 9

```
atg gtc ctc ctg agc atc ctt aga atc ctt ctt tgg gga ctg gtg ctt       48
Met Val Leu Leu Ser Ile Leu Arg Ile Leu Leu Trp Gly Leu Val Leu
1               5                  10                  15

ttt atg gaa cat agg gtc caa atg aca cag gta ggg cag ccc tct att       96
Phe Met Glu His Arg Val Gln Met Thr Gln Val Gly Gln Pro Ser Ile
            20                  25                  30

gcc cac ctg cct gag gcc cct acc ttg ccc ctg att cag gag ctg cta      144
Ala His Leu Pro Glu Ala Pro Thr Leu Pro Leu Ile Gln Glu Leu Leu
        35                  40                  45

gaa gaa gcc cct ggc aag cag cag agg aag ccg cgg gtc tta ggg cat      192
Glu Glu Ala Pro Gly Lys Gln Gln Arg Lys Pro Arg Val Leu Gly His
    50                  55                  60

ccc tta cgg tat atg ctg gag ctg tac cag cgt tca gct gac gca agt      240
Pro Leu Arg Tyr Met Leu Glu Leu Tyr Gln Arg Ser Ala Asp Ala Ser
65                  70                  75                  80

gga cac cct agg gaa aac cgc acc att ggg gcc acc atg gtg agg ctg      288
Gly His Pro Arg Glu Asn Arg Thr Ile Gly Ala Thr Met Val Arg Leu
                85                  90                  95
```

```
g agg ccg ctg gct agt gta gca agg cct ctc aga ggc tcc tgg cac      336
  Arg Pro Leu Ala Ser Val Ala Arg Pro Leu Arg Gly Ser Trp His
      100                 105                 110

ata cag acc ctg gac ttt cct ctg aga cca aac cgg gta gca tac caa      384
Ile Gln Thr Leu Asp Phe Pro Leu Arg Pro Asn Arg Val Ala Tyr Gln
    115                 120                 125

cta gtc aga gcc act gtg gtt tac cgc cat cag ctt cac cta act cat      432
Leu Val Arg Ala Thr Val Val Tyr Arg His Gln Leu His Leu Thr His
130                 135                 140

tcc cac ctc tcc tgc cat gtg gag ccc tgg gtc cag aaa agc cca acc      480
Ser His Leu Ser Cys His Val Glu Pro Trp Val Gln Lys Ser Pro Thr
145                 150                 155                 160

aat cac ttt cct tct tca gga aga ggc tcc tca aag cct tcc ctg ttg      528
Asn His Phe Pro Ser Ser Gly Arg Gly Ser Ser Lys Pro Ser Leu Leu
                165                 170                 175

ccc aaa act tgg aca gag atg gat atc atg gaa cat gtt ggg caa aag      576
Pro Lys Thr Trp Thr Glu Met Asp Ile Met Glu His Val Gly Gln Lys
            180                 185                 190

ctc tgg aat cac aag ggg cgc agg gtt cta cga ctc cgc ttc gtg tgt      624
Leu Trp Asn His Lys Gly Arg Arg Val Leu Arg Leu Arg Phe Val Cys
        195                 200                 205

cag cag cca aga ggt agt gag gtt ctt gag ttc tgg tgg cat ggc act      672
Gln Gln Pro Arg Gly Ser Glu Val Leu Glu Phe Trp Trp His Gly Thr
    210                 215                 220

tca tca ttg gac act gtc ttc ttg tta ctg tat ttc aat gac act          717
Ser Ser Leu Asp Thr Val Phe Leu Leu Leu Tyr Phe Asn Asp Thr
225                 230                 235

tagagtgttc agaagaccaa acctctccct aaaggcctga aagagtttac agaaaaagac    777

ccttctcttc tcttgaggag ggctcgtcaa gcaggcagta ttgcatcgga agttcctggc    837

ccctccaggg agcatgatgg gcctgaaagt aaccagtgtt ccctccaccc ttttcaagtc    897

agcttccagc agctgggctg ggatcactgg atcattgctc cccatctcta taccccaaac    957

tactgtaagg gagtatgtcc tcgggtacta cactatggtc tcaattctcc caatcatgcc   1017

atcatccaga accttgtcag tgagctggtg gatcagaatg tccctcagcc ttcctgtgtc   1077

ccttataagt atgttcccat tagcatcctt ctgattgagg caaatgggag tatcttgtac   1137

aaggagtatg agggtatgat tgcccagtcc tgcacatgca ggtga                   1182
```

<210> 10

<211> 239

<212> PRT

<213> Ovis areis

<220>

<221> misc_feature

<222> (1)..(3)

<223> atg start codon.
<220>
<221> misc_feature
<222> (805)..(807)
<223> first codon of mature peptide in wildtype sheep.
<220>
<221> misc_feature
<222> (1180)..(1182)
<223> tga stop codon.
<400> 10

```
Met Val Leu Leu Ser Ile Leu Arg Ile Leu Leu Trp Gly Leu Val Leu
1               5                  10                  15

Phe Met Glu His Arg Val Gln Met Thr Gln Val Gly Gln Pro Ser Ile
            20                  25                  30

Ala His Leu Pro Glu Ala Pro Thr Leu Pro Leu Ile Gln Glu Leu Leu
        35                  40                  45

Glu Glu Ala Pro Gly Lys Gln Gln Arg Lys Pro Arg Val Leu Gly His
    50                  55                  60

Pro Leu Arg Tyr Met Leu Glu Leu Tyr Gln Arg Ser Ala Asp Ala Ser
65                  70                  75                  80

Gly His Pro Arg Glu Asn Arg Thr Ile Gly Ala Thr Met Val Arg Leu
                85                  90                  95

Val Arg Pro Leu Ala Ser Val Ala Arg Pro Leu Arg Gly Ser Trp His
            100                 105                 110

Ile Gln Thr Leu Asp Phe Pro Leu Arg Pro Asn Arg Val Ala Tyr Gln
        115                 120                 125

Leu Val Arg Ala Thr Val Val Tyr Arg His Gln Leu His Leu Thr His
    130                 135                 140

Ser His Leu Ser Cys His Val Glu Pro Trp Val Gln Lys Ser Pro Thr
145                 150                 155                 160

Asn His Phe Pro Ser Ser Gly Arg Gly Ser Ser Lys Pro Ser Leu Leu
                165                 170                 175

Pro Lys Thr Trp Thr Glu Met Asp Ile Met Glu His Val Gly Gln Lys
            180                 185                 190
```

Trp Asn His Lys Gly Arg Arg Val Leu Arg Leu Arg Phe Val Cys
195 200 205

Gln Gln Pro Arg Gly Ser Glu Val Leu Glu Phe Trp Trp His Gly Thr
210 215 220

Ser Ser Leu Asp Thr Val Phe Leu Leu Leu Tyr Phe Asn Asp Thr
225 230 235

<210> 11

<211> 168

<212> DNA

<213> Ovis aries

<220>

<221> CDS

<222> (1)..(84)

<220>

<221> mutation

<222> (85)..(87)

<223> c to t at 85 of [S1] sheep changes glutamine cag codon to tag STOP

<400> 11
aga ggt agt gag gtt ctt gag ttc tgg tgg cat ggc act tca tca ttg 48
Arg Gly Ser Glu Val Leu Glu Phe Trp Trp His Gly Thr Ser Ser Leu
1 5 10 15

gac act gtc ttc ttg tta ctg tat ttc aat gac act tagagtgttc 94
Asp Thr Val Phe Leu Leu Leu Tyr Phe Asn Asp Thr
20 25

agaagaccaa acctctccct aaaggcctga aagagtttac agaaaaagac ccttctcttc 154

tcttgaggag ggct 168

<210> 12

<211> 28

<212> PRT

<213> Ovis aries

<400> 12

Gly Ser Glu Val Leu Glu Phe Trp Trp His Gly Thr Ser Ser Leu
5 10 15

Asp Thr Val Phe Leu Leu Leu Tyr Phe Asn Asp Thr
20 25

<210> 13

<211> 1665

<212> DNA

<213> Ovis aries

<220>

<221> 5'UTR

<222> (1)..(252)

<220>

<221> misc_feature

<222> (253)..(255)

<223> atg start codon.

<220>

<221> CDS

<222> (253)..(577)

<220>

<221> CDS

<222> (774)..(1627)

<220>

<221> Intron

<222> (578)..(773)

<223> n at 685 represents approx 5.2 kb intron.

<220>

<221> mat_peptide

<222> (1253)..()

<220>

<221> misc_feature

<222> (685)..()

<223> n represents approx 5.2 kb of intron

<220>

<221> misc_feature

<222> (1628)..(1630)

<223> tga stop codon.

<220>

<221> 3'UTR

<222> (1628)..(1665)

<220>

<221> mutation

<222> (1547)..(1549)

<223> g to t at 1548 of [S2] sheep changes agc serine codon to atc isoleucine codon

<400> 13

catgctgcct tgtcccacct gctgtttctg tttgtttgat gcaaagagga caatttagaa 60

gacctctttt tggttcagga gatcctacca gaggaagaaa cataggacct gcctgccagc 120

ctttcatttt tccttgccct atcctttgtg gtagtggagc ctggatgctg ttacccatgt 180

aaaaggaaag gtttaaagcg ttatcctttg ggcttttatc agaacatgtt gctgaacacc 240

aagcttttca ag atg gtc ctc ctg agc atc ctt aga atc ctt ctt tgg 288
Met Val Leu Leu Ser Ile Leu Arg Ile Leu Leu Trp
-265 -260

gga ctg gtg ctt ttt atg gaa cat agg gtc caa atg aca cag gta 333
Gly Leu Val Leu Phe Met Glu His Arg Val Gln Met Thr Gln Val
-255 -250 -245

ggg cag ccc tct att gcc cac ctg cct gag gcc cct acc ttg ccc 378
Gly Gln Pro Ser Ile Ala His Leu Pro Glu Ala Pro Thr Leu Pro
-240 -235 -230

ctg att cag gag ctg cta gaa gaa gcc cct ggc aag cag cag agg 423
Leu Ile Gln Glu Leu Leu Glu Glu Ala Pro Gly Lys Gln Gln Arg
-225 -220 -215

aag ccg cgg gtc tta ggg cat ccc tta cgg tat atg ctg gag ctg 468
Lys Pro Arg Val Leu Gly His Pro Leu Arg Tyr Met Leu Glu Leu

```
     -210                  -205                  -200

   c cag  cgt tca gct gac gca  agt gga cac cct agg  gaa aac cgc      513
Tyr Gln  Arg Ser Ala Asp Ala  Ser Gly His Pro Arg  Glu Asn Arg
    -195                  -190                  -185

acc att  ggg gcc acc atg gtg  agg ctg gtg agg ccg  ctg gct agt      558
Thr Ile  Gly Ala Thr Met Val  Arg Leu Val Arg Pro  Leu Ala Ser
    -180                  -175                  -170

gta gca  agg cct ctc aga g gtgagttatc atactatatt gttctggtgg       607
Val Ala  Arg Pro Leu Arg
    -165

gagggggga gaaaatgggg aagaaaagtg tagaaaaaag tggatctgtc agttttctgt  667

caggcttcac attgcctnca gtttgtactg agcaggtctg ttagagagac taaggctagg  727

atataagaag ctaacgcttt gctcttgttc cctcttacta atgcag gc   tcc tgg    781
                                                  Gly  Ser Trp
                                                  -160

    cac ata cag  acc ctg gac ttt cct  ctg aga cca aac cgg  gta gca     826
    His Ile Gln  Thr Leu Asp Phe Pro  Leu Arg Pro Asn Arg  Val Ala
            -155                  -150                  -145

    tac caa cta  gtc aga gcc act gtg  gtt tac cgc cat cag  ctt cac     871
    Tyr Gln Leu  Val Arg Ala Thr Val  Val Tyr Arg His Gln  Leu His
            -140                  -135                  -130

    cta act cat  tcc cac ctc tcc tgc  cat gtg gag ccc tgg  gtc cag     916
    Leu Thr His  Ser His Leu Ser Cys  His Val Glu Pro Trp  Val Gln
            -125                  -120                  -115

    aaa agc cca  acc aat cac ttt cct  tct tca gga aga ggc  tcc tca     961
    Lys Ser Pro  Thr Asn His Phe Pro  Ser Ser Gly Arg Gly  Ser Ser
            -110                  -105                  -100

    aag cct tcc ctg ttg ccc aaa act tgg aca gag atg gat atc atg gaa    1009
    Lys Pro Ser Leu Leu Pro Lys Thr Trp Thr Glu Met Asp Ile Met Glu
            -95                   -90                   -85

    cat gtt ggg caa aag ctc tgg aat cac aag ggg cgc agg gtt cta cga    1057
    His Val Gly Gln Lys Leu Trp Asn His Lys Gly Arg Arg Val Leu Arg
        -80                   -75                   -70

    ctc cgc ttc gtg tgt cag cag cca aga ggt agt gag gtt ctt gag ttc    1105
    Leu Arg Phe Val Cys Gln Gln Pro Arg Gly Ser Glu Val Leu Glu Phe
    -65                   -60                   -55                   -50

    tgg tgg cat ggc act tca tca ttg gac act gtc ttc ttg tta ctg tat    1153
    Trp Trp His Gly Thr Ser Ser Leu Asp Thr Val Phe Leu Leu Leu Tyr
                        -45                   -40                   -35

    ttc aat gac act cag agt gtt cag aag acc aaa cct ctc cct aaa ggc    1201
    Phe Asn Asp Thr Gln Ser Val Gln Lys Thr Lys Pro Leu Pro Lys Gly
                    -30                   -25                   -20

    ctg aaa gag ttt aca gaa aaa gac cct tct ctt ctc ttg agg agg gct    1249
    Leu Lys Glu Phe Thr Glu Lys Asp Pro Ser Leu Leu Leu Arg Arg Ala
            -15                   -10                   -5

    cgt caa gca ggc agt att gca tcg gaa gtt cct ggc ccc tcc agg gag    1297
    Arg Gln Ala Gly Ser Ile Ala Ser Glu Val Pro Gly Pro Ser Arg Glu
    -1  1                 5                   10                  15

    cat gat ggg cct gaa agt aac cag tgt tcc ctc cac cct ttt caa gtc    1345
    His Asp Gly Pro Glu Ser Asn Gln Cys Ser Leu His Pro Phe Gln Val
```

```
                20                  25                  30
ggc ttc cag cag ctg ggc tgg gat cac tgg atc att gct ccc cat ctc      1393
Ser Phe Gln Gln Leu Gly Trp Asp His Trp Ile Ile Ala Pro His Leu
            35                  40                  45

tat acc cca aac tac tgt aag gga gta tgt cct cgg gta cta cac tat      1441
Tyr Thr Pro Asn Tyr Cys Lys Gly Val Cys Pro Arg Val Leu His Tyr
        50                  55                  60

ggt ctc aat tct ccc aat cat gcc atc atc cag aac ctt gtc agt gag      1489
Gly Leu Asn Ser Pro Asn His Ala Ile Ile Gln Asn Leu Val Ser Glu
    65                  70                  75

ctg gtg gat cag aat gtc cct cag cct tcc tgt gtc cct tat aag tat      1537
Leu Val Asp Gln Asn Val Pro Gln Pro Ser Cys Val Pro Tyr Lys Tyr
80                  85                  90                  95

gtt ccc att atc atc ctt ctg att gag gca aat ggg agt atc ttg tac      1585
Val Pro Ile Ile Ile Leu Leu Ile Glu Ala Asn Gly Ser Ile Leu Tyr
                100                 105                 110

aag gag tat gag ggt atg att gcc cag tcc tgc aca tgc agg              1627
Lys Glu Tyr Glu Gly Met Ile Ala Gln Ser Cys Thr Cys Arg
            115                 120                 125

tgacggcaaa ggtgcagcta gctcaggttt cccaagaa                            1665
```

<210> 14

<211> 393

<212> PRT

<213> Ovis aries

<220>

<221> misc_feature

<222> (253)..(255)

<223> atg start codon.

<220>

<221> misc_feature

<222> (685)..()

<223> n represents approx 5.2 kb of intron

<220>

<221> misc_feature

<222> (1628)..(1630)

<223> tga stop codon.

<400> 14

```
Met Val Leu Leu  Ser Ile Leu Arg Ile  Leu Leu Trp Gly Leu  Val
            -265                -260                -255
```

```
Phe Met Glu His Arg Val Gln Met Thr Gln Val Gly Gln Pro
           -250                -245                -240

Ser Ile Ala His Leu Pro Glu Ala Pro Thr Leu Pro Leu Ile Gln
           -235                -230                -225

Glu Leu Leu Glu Glu Ala Pro Gly Lys Gln Gln Arg Lys Pro Arg
           -220                -215                -210

Val Leu Gly His Pro Leu Arg Tyr Met Leu Glu Leu Tyr Gln Arg
           -205                -200                -195

Ser Ala Asp Ala Ser Gly His Pro Arg Glu Asn Arg Thr Ile Gly
           -190                -185                -180

Ala Thr Met Val Arg Leu Val Arg Pro Leu Ala Ser Val Ala Arg
           -175                -170                -165

Pro Leu Arg Gly Ser Trp His Ile Gln Thr Leu Asp Phe Pro Leu
           -160                -155                -150

Arg Pro Asn Arg Val Ala Tyr Gln Leu Val Arg Ala Thr Val Val
           -145                -140                -135

Tyr Arg His Gln Leu His Leu Thr His Ser His Leu Ser Cys His
           -130                -125                -120

Val Glu Pro Trp Val Gln Lys Ser Pro Thr Asn His Phe Pro Ser
           -115                -110                -105

Ser Gly Arg Gly Ser Ser Lys Pro Ser Leu Leu Pro Lys Thr Trp Thr
           -100                -95                 -90

Glu Met Asp Ile Met Glu His Val Gly Gln Lys Leu Trp Asn His Lys
           -85                 -80                 -75

Gly Arg Arg Val Leu Arg Leu Arg Phe Val Cys Gln Gln Pro Arg Gly
     -70                 -65                 -60

Ser Glu Val Leu Glu Phe Trp Trp His Gly Thr Ser Ser Leu Asp Thr
-55                 -50                 -45                 -40

Val Phe Leu Leu Leu Tyr Phe Asn Asp Thr Gln Ser Val Gln Lys Thr
                -35                 -30                 -25

Lys Pro Leu Pro Lys Gly Leu Lys Glu Phe Thr Glu Lys Asp Pro Ser
           -20                 -15                 -10

Leu Leu Leu Arg Arg Ala Arg Gln Ala Gly Ser Ile Ala Ser Glu Val
      -5                  -1  1                   5
```

Gly Pro Ser Arg Glu His Asp Gly Pro Glu Ser Asn Gln Cys Ser
15 20 25

Leu His Pro Phe Gln Val Ser Phe Gln Gln Leu Gly Trp Asp His Trp
30 35 40

Ile Ile Ala Pro His Leu Tyr Thr Pro Asn Tyr Cys Lys Gly Val Cys
45 50 55

Pro Arg Val Leu His Tyr Gly Leu Asn Ser Pro Asn His Ala Ile Ile
60 65 70

Gln Asn Leu Val Ser Glu Leu Val Asp Gln Asn Val Pro Gln Pro Ser
75 80 85

Cys Val Pro Tyr Lys Tyr Val Pro Ile Ile Ile Leu Leu Ile Glu Ala
90 95 100 105

Asn Gly Ser Ile Leu Tyr Lys Glu Tyr Glu Gly Met Ile Ala Gln Ser
110 115 120

Cys Thr Cys Arg
125

<210> 15

<211> 1182

<212> DNA

<213> Ovis aries

<220>

<221> misc_feature

<222> (1)..(3)

<223> atg start codon.

<220>

<221> mutation

<222> (1099)..(1101)

<223> g to to at 1100 of [S2] sheep changes agc serine codon to atc iso leucine codon

<220>

<221> CDS

<2> (1)..(1179)

<220>
<221> mat_peptide
<222> (805)..()

<220>
<221> misc_feature
<222> (1180)..(1182)
<223> tga stop codon.

<400> 15

```
atg gtc ctc ctg  agc atc ctt aga atc  ctt ctt tgg gga ctg  gtg        45
Met Val Leu Leu  Ser Ile Leu Arg Ile  Leu Leu Trp Gly Leu  Val
               -265                 -260                 -255

ctt ttt atg gaa  cat agg gtc caa atg  aca cag gta ggg cag  ccc        90
Leu Phe Met Glu  His Arg Val Gln Met  Thr Gln Val Gly Gln  Pro
               -250                 -245                 -240

tct att gcc cac  ctg cct gag gcc cct  acc ttg ccc ctg att  cag       135
Ser Ile Ala His  Leu Pro Glu Ala Pro  Thr Leu Pro Leu Ile  Gln
               -235                 -230                 -225

gag ctg cta gaa  gaa gcc cct ggc aag  cag cag agg aag ccg  cgg       180
Glu Leu Leu Glu  Glu Ala Pro Gly Lys  Gln Gln Arg Lys Pro  Arg
               -220                 -215                 -210

gtc tta ggg cat  ccc tta cgg tat atg  ctg gag ctg tac cag  cgt       225
Val Leu Gly His  Pro Leu Arg Tyr Met  Leu Glu Leu Tyr Gln  Arg
               -205                 -200                 -195

tca gct gac gca  agt gga cac cct agg  gaa aac cgc acc att  ggg       270
Ser Ala Asp Ala  Ser Gly His Pro Arg  Glu Asn Arg Thr Ile  Gly
               -190                 -185                 -180

gcc acc atg gtg  agg ctg gtg agg ccg  ctg gct agt gta gca  agg       315
Ala Thr Met Val  Arg Leu Val Arg Pro  Leu Ala Ser Val Ala  Arg
               -175                 -170                 -165

cct ctc aga ggc  tcc tgg cac ata cag  acc ctg gac ttt cct  ctg       360
Pro Leu Arg Gly  Ser Trp His Ile Gln  Thr Leu Asp Phe Pro  Leu
               -160                 -155                 -150

aga cca aac cgg  gta gca tac caa cta  gtc aga gcc act gtg  gtt       405
Arg Pro Asn Arg  Val Ala Tyr Gln Leu  Val Arg Ala Thr Val  Val
               -145                 -140                 -135

tac cgc cat cag  ctt cac cta act cat  tcc cac ctc tcc tgc  cat       450
Tyr Arg His Gln  Leu His Leu Thr His  Ser His Leu Ser Cys  His
               -130                 -125                 -120

gtg gag ccc tgg  gtc cag aaa agc cca  acc aat cac ttt cct  tct       495
Val Glu Pro Trp  Val Gln Lys Ser Pro  Thr Asn His Phe Pro  Ser
               -115                 -110                 -105

tca gga aga ggc  tcc tca aag cct tcc ctg ttg ccc aaa act tgg aca    543
```

```
Gly Arg Gly  Ser Ser Lys Pro Ser Leu Leu Pro Lys Thr Trp Thr
        -100                -95                 -90

gag atg gat atc atg gaa cat gtt ggg caa aag ctc tgg aat cac aag      591
Glu Met Asp Ile Met Glu His Val Gly Gln Lys Leu Trp Asn His Lys
     -85                 -80                 -75

ggg cgc agg gtt cta cga ctc cgc ttc gtg tgt cag cag cca aga ggt      639
Gly Arg Arg Val Leu Arg Leu Arg Phe Val Cys Gln Gln Pro Arg Gly
 -70                 -65                 -60

agt gag gtt ctt gag ttc tgg tgg cat ggc act tca tca ttg gac act      687
Ser Glu Val Leu Glu Phe Trp Trp His Gly Thr Ser Ser Leu Asp Thr
-55                 -50                 -45                 -40

gtc ttc ttg tta ctg tat ttc aat gac act cag agt gtt cag aag acc      735
Val Phe Leu Leu Leu Tyr Phe Asn Asp Thr Gln Ser Val Gln Lys Thr
            -35                 -30                 -25

aaa cct ctc cct aaa ggc ctg aaa gag ttt aca gaa aaa gac cct tct      783
Lys Pro Leu Pro Lys Gly Leu Lys Glu Phe Thr Glu Lys Asp Pro Ser
        -20                 -15                 -10

ctt ctc ttg agg agg gct cgt caa gca ggc agt att gca tcg gaa gtt      831
Leu Leu Leu Arg Arg Ala Arg Gln Ala Gly Ser Ile Ala Ser Glu Val
    -5                  -1  1                   5

cct ggc ccc tcc agg gag cat gat ggg cct gaa agt aac cag tgt tcc      879
Pro Gly Pro Ser Arg Glu His Asp Gly Pro Glu Ser Asn Gln Cys Ser
10                  15                  20                  25

ctc cac cct ttt caa gtc agc ttc cag cag ctg ggc tgg gat cac tgg      927
Leu His Pro Phe Gln Val Ser Phe Gln Gln Leu Gly Trp Asp His Trp
                30                  35                  40

atc att gct ccc cat ctc tat acc cca aac tac tgt aag gga gta tgt      975
Ile Ile Ala Pro His Leu Tyr Thr Pro Asn Tyr Cys Lys Gly Val Cys
            45                  50                  55

cct cgg gta cta cac tat ggt ctc aat tct ccc aat cat gcc atc atc     1023
Pro Arg Val Leu His Tyr Gly Leu Asn Ser Pro Asn His Ala Ile Ile
        60                  65                  70

cag aac ctt gtc agt gag ctg gtg gat cag aat gtc cct cag cct tcc     1071
Gln Asn Leu Val Ser Glu Leu Val Asp Gln Asn Val Pro Gln Pro Ser
    75                  80                  85

tgt gtc cct tat aag tat gtt ccc att atc atc ctt ctg att gag gca     1119
Cys Val Pro Tyr Lys Tyr Val Pro Ile Ile Ile Leu Leu Ile Glu Ala
90                  95                  100                 105

aat ggg agt atc ttg tac aag gag tat gag ggt atg att gcc cag tcc     1167
Asn Gly Ser Ile Leu Tyr Lys Glu Tyr Glu Gly Met Ile Ala Gln Ser
                110                 115                 120

tgc aca tgc agg tga                                                1182
Cys Thr Cys Arg
            125
```

<210> 16

<211> 393

<212> PRT

<213> Ovis aries

<220>

<221> misc_feature

<222> (1)..(3)

<223> atg start codon.

<220>

<221> misc_feature

<222> (1180)..(1182)

<223> tga stop codon.

<400> 16

```
Met Val Leu Leu  Ser Ile Leu Arg Ile  Leu Leu Trp Gly Leu  Val
               -265                 -260                 -255

Leu Phe Met Glu  His Arg Val Gln Met  Thr Gln Val Gly Gln  Pro
               -250                 -245                 -240

Ser Ile Ala His  Leu Pro Glu Ala Pro  Thr Leu Pro Leu Ile  Gln
               -235                 -230                 -225

Glu Leu Leu Glu  Glu Ala Pro Gly Lys  Gln Gln Arg Lys Pro  Arg
               -220                 -215                 -210

Val Leu Gly His  Pro Leu Arg Tyr Met  Leu Glu Leu Tyr Gln  Arg
               -205                 -200                 -195

Ser Ala Asp Ala  Ser Gly His Pro Arg  Glu Asn Arg Thr Ile  Gly
               -190                 -185                 -180

Ala Thr Met Val  Arg Leu Val Arg Pro  Leu Ala Ser Val Ala  Arg
               -175                 -170                 -165

Pro Leu Arg Gly  Ser Trp His Ile Gln  Thr Leu Asp Phe Pro  Leu
               -160                 -155                 -150

Arg Pro Asn Arg  Val Ala Tyr Gln Leu  Val Arg Ala Thr Val  Val
               -145                 -140                 -135

Tyr Arg His Gln  Leu His Leu Thr His  Ser His Leu Ser Cys  His
               -130                 -125                 -120

Val Glu Pro Trp  Val Gln Lys Ser Pro  Thr Asn His Phe Pro  Ser
               -115                 -110                 -105

Ser Gly Arg Gly  Ser Ser Lys Pro Ser  Leu Leu Pro Lys Thr  Trp  Thr
               -100                 -95                  -90
```

Met Asp Ile Met Glu His Val Gly Gln Lys Leu Trp Asn His Lys
-85 -80 -75

Gly Arg Arg Val Leu Arg Leu Arg Phe Val Cys Gln Gln Pro Arg Gly
-70 -65 -60

Ser Glu Val Leu Glu Phe Trp Trp His Gly Thr Ser Ser Leu Asp Thr
-55 -50 -45 -40

Val Phe Leu Leu Leu Tyr Phe Asn Asp Thr Gln Ser Val Gln Lys Thr
-35 -30 -25

Lys Pro Leu Pro Lys Gly Leu Lys Glu Phe Thr Glu Lys Asp Pro Ser
-20 -15 -10

Leu Leu Leu Arg Arg Ala Arg Gln Ala Gly Ser Ile Ala Ser Glu Val
-5 -1 1 5

Pro Gly Pro Ser Arg Glu His Asp Gly Pro Glu Ser Asn Gln Cys Ser
10 15 20 25

Leu His Pro Phe Gln Val Ser Phe Gln Gln Leu Gly Trp Asp His Trp
30 35 40

Ile Ile Ala Pro His Leu Tyr Thr Pro Asn Tyr Cys Lys Gly Val Cys
45 50 55

Pro Arg Val Leu His Tyr Gly Leu Asn Ser Pro Asn His Ala Ile Ile
60 65 70

Gln Asn Leu Val Ser Glu Leu Val Asp Gln Asn Val Pro Gln Pro Ser
75 80 85

Cys Val Pro Tyr Lys Tyr Val Pro Ile Ile Ile Leu Leu Ile Glu Ala
90 95 100 105

Asn Gly Ser Ile Leu Tyr Lys Glu Tyr Glu Gly Met Ile Ala Gln Ser
110 115 120

Cys Thr Cys Arg
125

<210> 17

<211> 168

<212> DNA

<213> Ovis aries

<220>

<221> CDS

<222> (1)..(168)

<220>

<221> mutation

<222> (85)..(87)

<223> g to t at 86 of GDF9B sheep changes agc serine codon to atc isoleucine codon

<400> 17

```
gcc atc atc cag aac ctt gtc agt gag ctg gtg gat cag aat gtc cct       48
Ala Ile Ile Gln Asn Leu Val Ser Glu Leu Val Asp Gln Asn Val Pro
1               5                   10                  15

cag cct tcc tgt gtc cct tat aag tat gtt ccc att atc atc ctt ctg       96
Gln Pro Ser Cys Val Pro Tyr Lys Tyr Val Pro Ile Ile Ile Leu Leu
            20                  25                  30

att gag gca aat ggg agt atc ttg tac aag gag tat gag ggt atg att      144
Ile Glu Ala Asn Gly Ser Ile Leu Tyr Lys Glu Tyr Glu Gly Met Ile
        35                  40                  45

gcc cag tcc tgc aca tgc agg tga                                      168
Ala Gln Ser Cys Thr Cys Arg
    50                  55
```

<210> 18

<211> 55

<212> PRT

<213> Ovis aries

<400> 18

```
Ala Ile Ile Gln Asn Leu Val Ser Glu Leu Val Asp Gln Asn Val Pro
1               5                   10                  15

Gln Pro Ser Cys Val Pro Tyr Lys Tyr Val Pro Ile Ile Ile Leu Leu
            20                  25                  30

Ile Glu Ala Asn Gly Ser Ile Leu Tyr Lys Glu Tyr Glu Gly Met Ile
        35                  40                  45

Ala Gln Ser Cys Thr Cys Arg
    50                  55
```

FIGURE 1: Irish Cambridge and F700 Belclare sheep GDF9 sequence and mutations.

B. Nucleotide substitutions of the four GDF9 mutations which change an amino acid compared with wild-type sheep sequence.

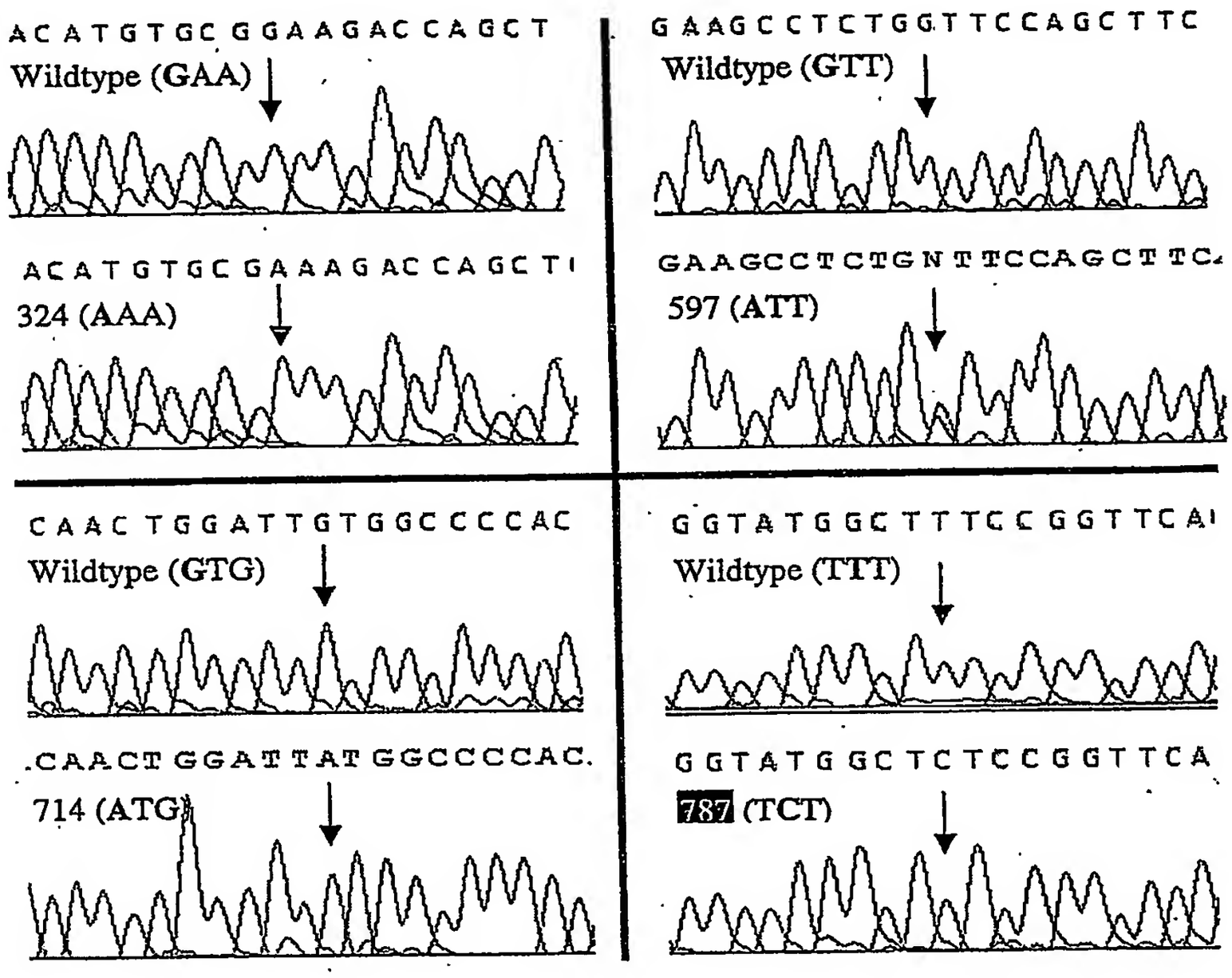


FIGURE 1: Irish Cambridge and F700 Belclare sheep GDF9 sequence and mutations.

A. Predicted amino acid sequence of sheep GDF9 protein. Numbers at the start of each line indicate amino acid positions of full-length unprocessed protein. Numbers in brackets indicate amino acid positions of the mature peptide. The RRHR furin protease cleavage site and predicted start of the mature processed peptide is shaded in grey. The filled triangle indicates the position of the single intron within the gene. The open triangles indicate positions of mutations that confer amino acid substitutions but are not associated with the sterility phenotype. The position of the [787] mutation associated with sterility is shaded black.

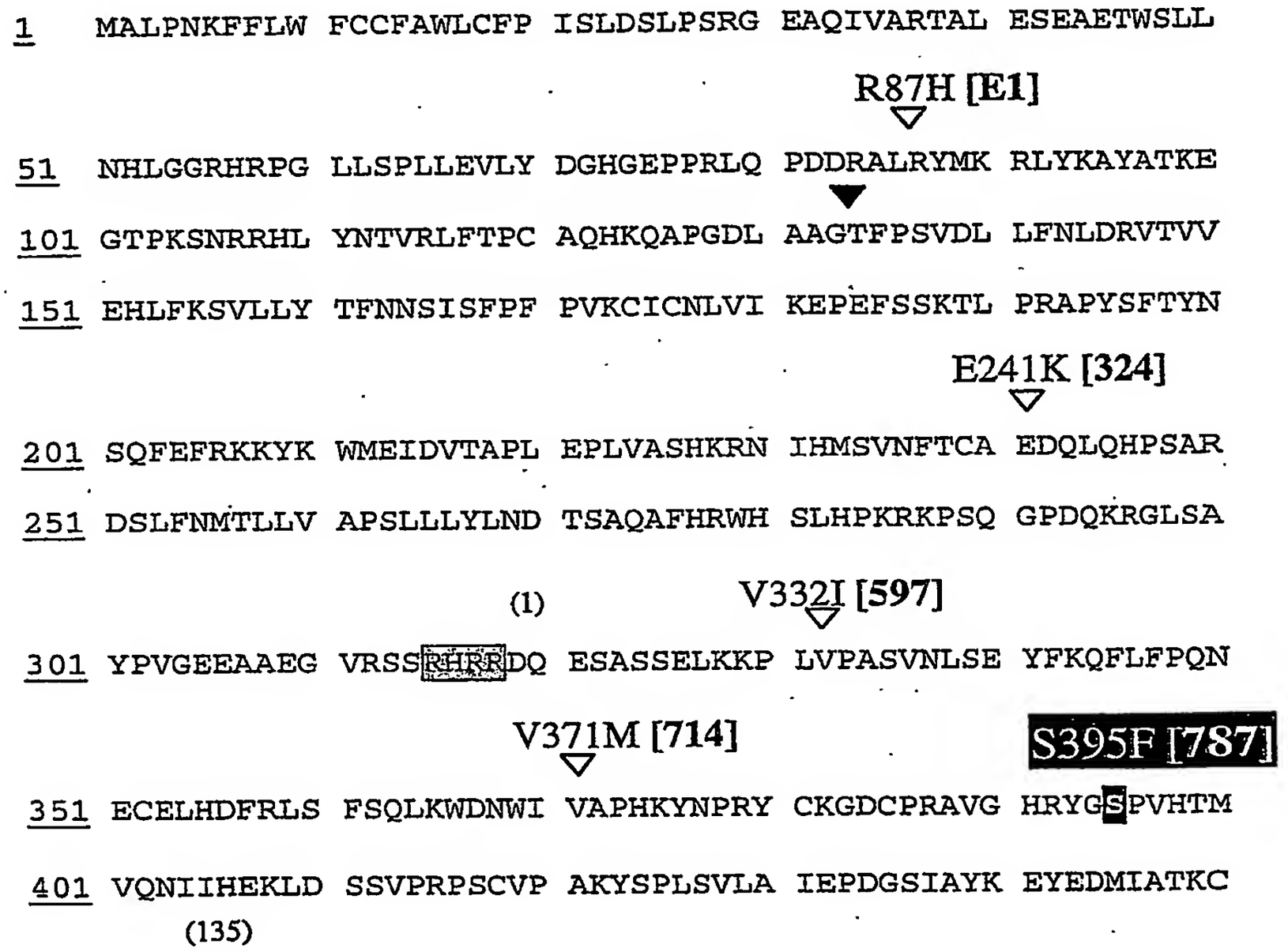


FIGURE 2: Irish Cambridge and F700 Belclare sheep GDF9B sequence and mutations.

B. Nucleotide substitutions of the two GDF9B mutations which change an amino acid compared with wild-type sheep sequence.

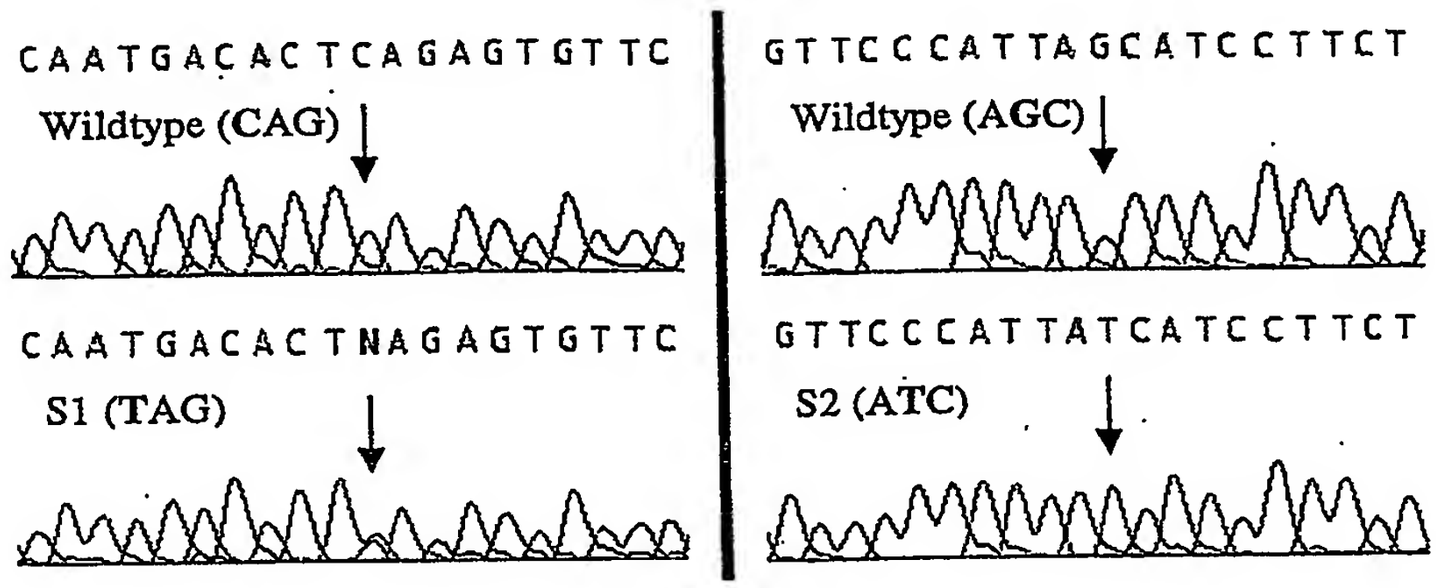


FIGURE 2: Irish Cambridge and F700 Belclare sheep GDF9B sequence and mutations.

a. Predicted amino acid sequence of sheep GDF9B protein. Numbers at the start of each line indicate amino acid positions of full-length unprocessed protein. Numbers in brackets indicate amino acid positions of the mature peptide. The RRAR furin protease cleavage site and predicted start of the mature processed peptide is shaded in grey. The filled triangle indicates the position of the single intron within the gene. The open triangle indicates the position of a single Leu deletion polymorphism. The position of the [S1] and [S2] mutations associated with sterility are shaded black.

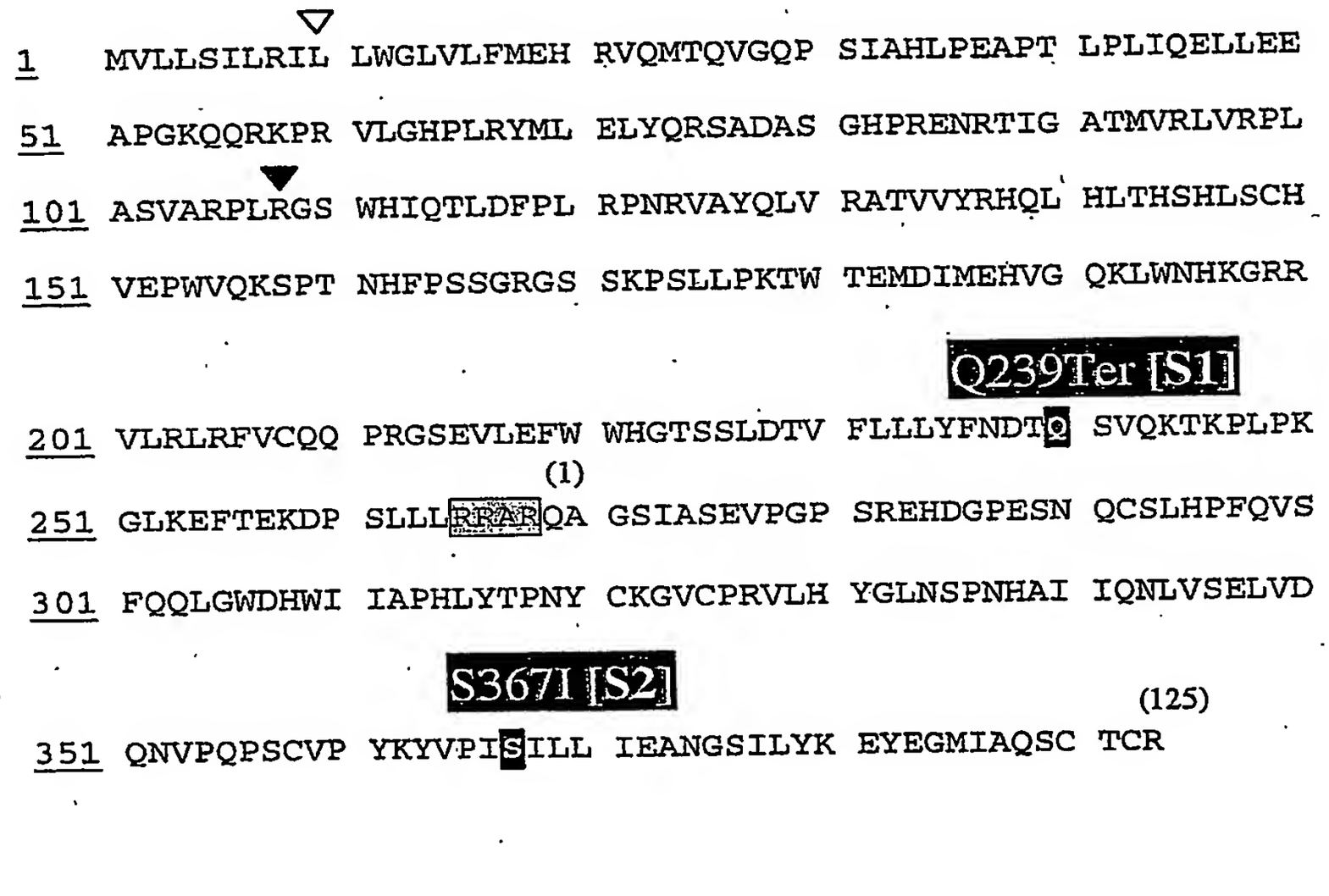
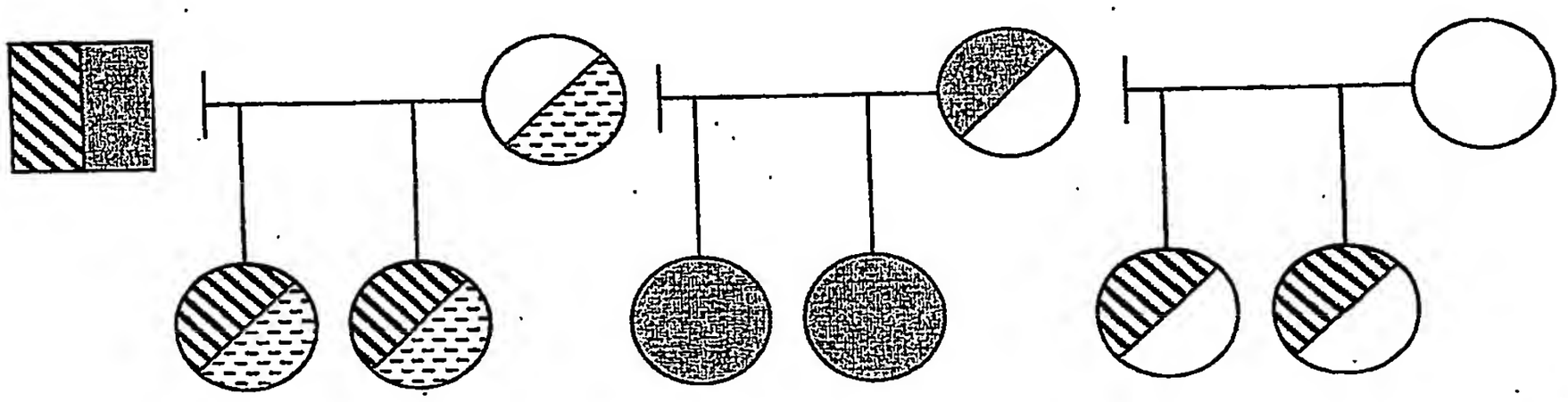


FIGURE 3: Schematic representation of genotypes within F700 Belclare and Cambridge pedigrees

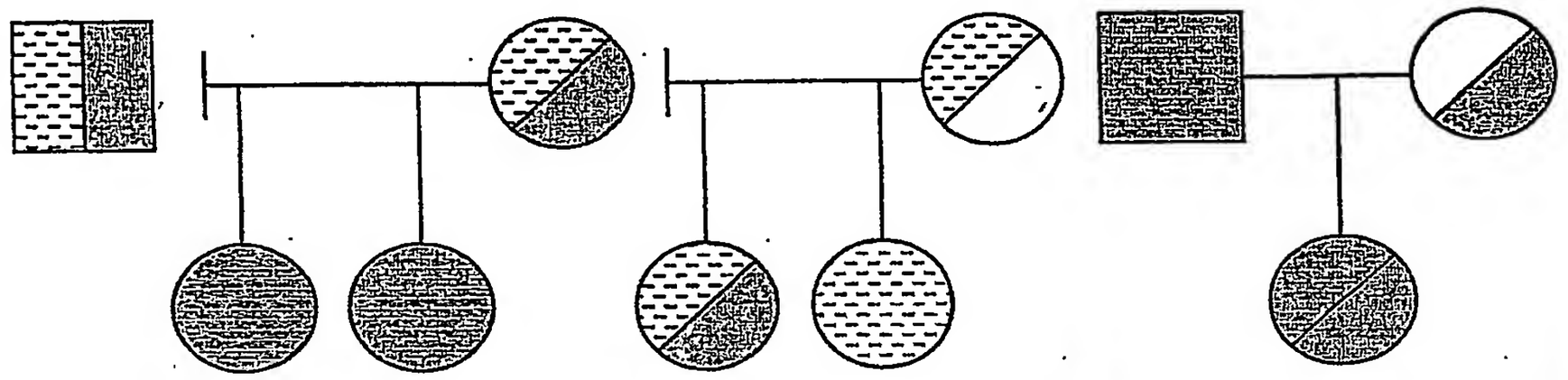
a. F700 Belclare pedigree. The pedigree represents sire R830 mated to three ewes 9704, 8783 and 7810, and their six female offspring. The table below the pedigree diagram shows the phenotype and genotypes for each of the animals. The phenotype of each animal is indicated as sterile S, fertile F or n/a (not applicable) for the male. Dashed lines within the circles (females) or squares (males) denote carriers of the GDF9B [S1] mutation, diagonal lines denote carriers of the GDF9B [S2] mutation and the solid grey denotes carriers of the GDF9 [787] mutation (T nucleotide). Half-cirles or half-squares denote those animals heterozygous for a mutation. Sire R830 is heterozygous for [787] (grey) but hemizygous for [S2] as GDF9B is carried on the X chromosome.



Animal		R830	930458	930459	9704	930811	930812	8783	930810	948302	7810
Phenotype Sterile (S) Fertile (F)		n/a	S	S	F	S	S	F	F	F	F
GENOTYPE	S1	+/Y	S1/+	S1/+	S1/+	+/+	+/+	+/+	+/+	+/+	+/+
	S2	S2/Y	S2/+	S2/+	+/+	S2/+	S2/+	+/+	S2/+	S2/+	+/+
	787	T/+	T/+	T/+	+/+	+/+	T/T	T/T	T/+	+/+	+/+

FIGURE 3: Schematic representation of genotypes within F700 Belclare and Cambridge pedigrees

b. Cambridge pedigrees. The pedigree represents sire 962101 mated to two ewes 962158 and 976234, and their four female offspring, and sire 930142 mated to ewe 8874 and their one female offspring. The table below the pedigree diagram shows the phenotype and genotypes for each of the animals. The phenotype of each animal is indicated as sterile S, fertile F or n/a (not applicable) for the male. Dashed lines within the circles (females) or squares (males) denote carriers of the GDF9B [S1] mutation and the solid grey denotes carriers of the GDF9 [787] mutation (T nucleotide). Dashed grey squares and circles denote carriers of both mutations. Half-and-half shading denotes those animals heterozygous for the mutations. Sire 962101 is heterozygous for [787] (grey) but hemizygous for [S2] as GDF9B is carried on the X chromosome.



Animal	962101	976344	976355	962152	975522	975533	976234	930142	948093	8874
Phenotype Sterile (S) Fertile (F)	n/a	S	S	F	F	S	F	n/a	S	F
GENOTYPE S1	S1/Y	S1/S1	S1/S1	S1/+	S1/+	S1/S1	S1/+	S1/Y	S1/+	+/+
GENOTYPE S2	+/Y	+/+	+/+	+/+	+/+	+/+	+/+	+/+	+/+	+/+
GENOTYPE 787	T/+	T/T	T/T	T/+	T/+	T/+	+/+	T/T	T/T	T/+

FIGURE 4: Nucleotide and amino acid of wildtype sheep GDF9 showing positions of mutations in Irish Cambridge and F700 Belclare sheep

Numbers on the right indicate nucleotide position from the atg start codon. Numbers under each line indicate amino acid residue numbering from the start of the mature processed active peptide. Negative numbers indicate amino acids in the pro-region of the protein. The position of the single intron is marked by white boxes inserted into the sequence. Positions of the eight nucleotide polymorphisms are marked in bold within boxes, and named according to Table 1 using square brackets [ ]. The amino acids residues which are changed by the nucleotide polymorphisms are underlined. The taa stop codon indicates the end of the protein.

```
Sheep GDF9
atg gcg ctt ccc  aac aaa ttc ttc ctt  tgg ttt tgc tgc ttt  gcc        45
Met Ala Leu Pro  Asn Lys Phe Phe Leu  Trp Phe Cys Cys Phe  Ala
          -315                -310                -305

tgg ctc tgt ttt  cct att agc ctt gat  tct ctg cct tct agg  gga        90
Trp Leu Cys Phe  Pro Ile Ser Leu Asp  Ser Leu Pro Ser Arg  Gly
          -300                -295                -290

gaa gct cag att  gta gct agg act gcg  ttg gaa tct gag gct  gag       135
Glu Ala Gln Ile  Val Ala Arg Thr Ala  Leu Glu Ser Glu Ala  Glu
          -285                -280                -275

act tgg tcc ttg  ctg aac cat tta ggt  ggg aga cac aga cct  ggt       180
Thr Trp Ser Leu  Leu Asn His Leu Gly  Gly Arg His Arg Pro  Gly
          -270                -265                -260

ctc ctt tcc cct  ctc tta gag gtt ctg  tat gat ggg cac ggg  gaa       225
Leu Leu Ser Pro  Leu Leu Glu Val Leu  Tyr Asp Gly His Gly  Glu
          -255                -250                -245
                                          [E1]
ccc ccc agg ctg  cag cca gat gac aga  gct ttg cgc tac atg  aag       270
Pro Pro Arg Leu  Gln Pro Asp Asp Arg  Ala Leu Arg Tyr Met  Lys
          -240                -235                -230

agg ctc tat aag  gca tac gct acc aag  gag ggg acc cct aaa  tcc       315
Arg Leu Tyr Lys  Ala Tyr Ala Thr Lys  Glu Gly Thr Pro Lys  Ser
          -225                -220                -215

aac aga cgc cac  ctc tac aac act gtt  cgg ctc ttc acc ccc  tgt       360
Asn Arg Arg His  Leu Tyr Asn Thr Val  Arg Leu Phe Thr Pro  Cys
          -210                -205                -200
                                          Intron position
gct cag cac aag  cag gct cct ggg gac  ctg gcg gca g ga acc  ttt       405
Ala Gln His Lys  Gln Ala Pro Gly Asp  Leu Ala Ala G ly Thr  Phe
          -195                -190                -185

cca tca gtg gat  ctg ctg ttt aac ctg  gat cgt gtt act gtt  gtg       450
Pro Ser Val Asp  Leu Leu Phe Asn Leu  Asp Arg Val Thr Val  Val
          -180                -175                -170
                          [74]     [80]
gaa cat tta ttc  aag tca gtc ttg ctg  tat act ttc aac aac  tcc       495
Glu His Leu Phe  Lys Ser Val Leu Leu  Tyr Thr Phe Asn Asn  Ser
          -165                -160                -155

att tct ttt ccc  ttt cct gtt aaa tgt  ata tgc aac ctg gtg  ata       540
Ile Ser Phe Pro  Phe Pro Val Lys Cys  Ile Cys Asn Leu Val  Ile
          -150                -145                -140

aaa gag cca gag  ttt tct agc aag act  ctc cct aga gct cca  tac       585
Lys Glu Pro Glu  Phe Ser Ser Lys Thr  Leu Pro Arg Ala Pro  Tyr
          -135                -130                -125

tca ttt acc tat  aac tca cag ttt gaa  ttt aga aag aaa tac  aaa       630
```

```
Ser Phe Thr Tyr  Asn Ser Gln Phe Glu  Phe Arg Lys Lys Tyr  Lys
            -120                 -115                 -110

tgg atg gag att  gat gtg acg gct cct  ctt gag cct ctg gtg gcc tcc    678
Trp Met Glu Ile  Asp Val Thr Ala Pro  Leu Glu Pro Leu Val Ala Ser
            -105                 -100                 -95
                                                      [324]
cac aag agg aat att cac atg tct gta aat ttt aca tgt gcg gaa gac    726
His Lys Arg Asn Ile His Met Ser Val Asn Phe Thr Cys Ala Glu Asp
        -90                  -85                  -80

cag ctg cag cat cct tca gcg cgg gac agc ctg ttt aac atg act ctt    774
Gln Leu Gln His Pro Ser Ala Arg Asp Ser Leu Phe Asn Met Thr Leu
    -75                  -70                  -65

ctc gta gcg ccc tca ctg ctt ttg tat ctg aac gac aca agt gct cag    822
Leu Val Ala Pro Ser Leu Leu Leu Tyr Leu Asn Asp Thr Ser Ala Gln
-60                  -55                  -50                  -45

gct ttt cac agg tgg cat tcc ctc cac cct aaa agg aag cct tca cag    870
Ala Phe His Arg Trp His Ser Leu His Pro Lys Arg Lys Pro Ser Gln
                -40                  -35                  -30

ggt cct gac cag aag aga ggg cta tct gcc tac ccc gtg gga gaa gaa    918
Gly Pro Asp Gln Lys Arg Gly Leu Ser Ala Tyr Pro Val Gly Glu Glu
            -25                  -20                  -15

gct gct gag ggt gta aga tcg tcc cgt cac cgc aga gac cag gag agt    966
Ala Ala Glu Gly Val Arg Ser Ser Arg His Arg Arg Asp Gln Glu Ser
        -10                  -5                   -1  1
             [581]                      [597]
gcc agc tct gaa ttg aag aag cct ctg gtt cca gct tca gtc aat ctg   1014
Ala Ser Ser Glu Leu Lys Lys Pro Leu Val Pro Ala Ser Val Asn Leu
5                   10                  15                  20

agt gaa tac ttc aaa cag ttt ctt ttt ccc cag aat gaa tgt gag ctc   1062
Ser Glu Tyr Phe Lys Gln Phe Leu Phe Pro Gln Asn Glu Cys Glu Leu
                25                  30                  35

cat gac ttt aga ctt agc ttt agt cag ctg aag tgg gac aac tgg att   1110
His Asp Phe Arg Leu Ser Phe Ser Gln Leu Lys Trp Asp Asn Trp Ile
            40                  45                  50
[714]
gtg gcc cca cac aaa tac aac cct cga tac tgt aaa ggg gac tgt ccc   1158
Val Ala Pro His Lys Tyr Asn Pro Arg Tyr Cys Lys Gly Asp Cys Pro
        55                  60                  65
                                  [787]
agg gcg gtc gga cat cgg tat ggc tct ccg gtt cac acc atg gtg cag   1206
Arg Ala Val Gly His Arg Tyr Gly Ser Pro Val His Thr Met Val Gln
    70                  75                  80

aac atc atc cat gag aaa ctt gac tcc tca gtg cca aga cca tcc tgt   1254
Asn Ile Ile His Glu Lys Leu Asp Ser Ser Val Pro Arg Pro Ser Cys
85                  90                  95                  100

gta cct gcc aag tat agc cct ttg agt gtt ttg gcc atc gag cct gat   1302
Val Pro Ala Lys Tyr Ser Pro Leu Ser Val Leu Ala Ile Glu Pro Asp
                105                 110                 115

ggc tca atc gct tat aaa gaa tat gaa gat atg ata gcc act aag tgt   1350
Gly Ser Ile Ala Tyr Lys Glu Tyr Glu Asp Met Ile Ala Thr Lys Cys
            120                 125                 130

acc tgt cgt taa cagactc ctgtcaagta aaaccatgag tgtcctggcc          1399
Thr Cys Arg STOP
        135

agtgtaaatg ccgcgcc                                               1416
```

FIGURE 5: Nucleotide and amino acid of wildtype sheep GDF9B showing positions of mutations in Irish Cambridge and F700 Belclare sheep

Numbers on the right indicate nucleotide position from the atg start codon. Numbers under each line indicate amino acid residue numbering from the start of the mature processed active peptide. Negative numbers indicate amino acids in the pro-region of the protein. The position of the single intron is marked by a dashed inserted into the sequence. Positions of the four nucleotide polymorphisms are marked in bold within boxes, and named according to Table 1 using square brackets [ ]. The amino acids residues which are changed by the nucleotide polymorphisms are underlined. Asterisks (***) indicate the positions of the previously reported Hanna (Gln to Stop codon) and Inverdale (Val to Asp codon) mutations. The tga stop codon indicates the end of the protein.

Sheep BMP15 full

```
. . . . . .atg gtc ctc ctg  agc atc ctt aga atc        27
. . . . . .Met Val Leu Leu  Ser Ile Leu Arg Ile
                          -265                 -260

[Leu] deletion
ctt ctt tgg gga ctg  gtg ctt ttt atg gaa  cat agg gtc caa atg      72
Leu Leu Trp Gly Leu  Val Leu Phe Met Glu  His Arg Val Gln Met
               -255                 -250                 -245

aca cag gta ggg cag  ccc tct att gcc cac  ctg cct gag gcc cct      117
Thr Gln Val Gly Gln  Pro Ser Ile Ala His  Leu Pro Glu Ala Pro
               -240                 -235                 -230

acc ttg ccc ctg att  cag gag ctg cta gaa  gaa gcc cct ggc aag      162
Thr Leu Pro Leu Ile  Gln Glu Leu Leu Glu  Glu Ala Pro Gly Lys
               -225                 -220                 -215

cag cag agg aag ccg  cgg gtc tta ggg cat  ccc tta cgg tat atg      207
Gln Gln Arg Lys Pro  Arg Val Leu Gly His  Pro Leu Arg Tyr Met
               -210                 -205                 -200

ctg gag ctg tac cag  cgt tca gct gac gca  agt gga cac cct agg      252
Leu Glu Leu Tyr Gln  Arg Ser Ala Asp Ala  Ser Gly His Pro Arg
               -195                 -190                 -185

gaa aac cgc acc att  ggg gcc acc atg gtg  agg ctg gtg agg ccg      297
Glu Asn Arg Thr Ile  Gly Ala Thr Met Val  Arg Leu Val Arg Pro
               -180                 -175                 -170
                                  Intron position
ctg gct agt gta gca  agg cct ctc aga g----------------------gc      327
Leu Ala Ser Val Ala  Arg Pro Leu Arg                       Gly
               -165                                        -160

tcc tgg cac ata cag  acc ctg gac ttt cct  ctg aga cca aac cgg      372
Ser Trp His Ile Gln  Thr Leu Asp Phe Pro  Leu Arg Pro Asn Arg
               -155                 -150                 -145

gta gca tac caa cta  gtc aga gcc act gtg  gtt tac cgc cat cag      417
Val Ala Tyr Gln Leu  Val Arg Ala Thr Val  Val Tyr Arg His Gln
               -140                 -135                 -130

ctt cac cta act cat  tcc cac ctc tcc tgc  cat gtg gag ccc tgg      462
Leu His Leu Thr His  Ser His Leu Ser Cys  His Val Glu Pro Trp
               -125                 -120                 -115

gtc cag aaa agc cca  acc aat cac ttt cct  tct tca gga aga ggc      507
```

```
Val Gln Lys Ser Pro  Thr Asn His Phe Pro  Ser Ser Gly Arg Gly
                -110                -105                -100

tcc tca aag cct tcc ctg ttg ccc aaa act tgg aca gag atg gat atc      555
Ser Ser Lys Pro Ser Leu Leu Pro Lys Thr Trp Thr Glu Met Asp Ile
                -95                 -90                 -85

atg gaa cat gtt ggg caa aag ctc tgg aat cac aag ggg cgc agg gtt      603
Met Glu His Val Gly Gln Lys Leu Trp Asn His Lys Gly Arg Arg Val
            -80                 -75                 -70

cta cga ctc cgc ttc gtg tgt cag cag cca aga ggt agt gag gtt ctt      651
Leu Arg Leu Arg Phe Val Cys Gln Gln Pro Arg Gly Ser Glu Val Leu
        -65                 -60                 -55

gag ttc tgg tgg cat ggc act tca tca ttg gac act gtc ttc ttg tta      699
Glu Phe Trp Trp His Gly Thr Ser Ser Leu Asp Thr Val Phe Leu Leu
    -50                 -45                 -40

                        [S1]                                [422]
ctg tat ttc aat gac act cag agt gtt cag aag acc aaa cct ctc cct      747
Leu Tyr Phe Asn Asp Thr Gln Ser Val Gln Lys Thr Lys Pro Leu Pro
-35                 -30                 -25                 -20

aaa ggc ctg aaa gag ttt aca gaa aaa gac cct tct ctt ctc ttg agg      795
Lys Gly Leu Lys Glu Phe Thr Glu Lys Asp Pro Ser Leu Leu Leu Arg
                -15                 -10                 -5

agg gct cgt caa gca ggc agt att gca tcg gaa gtt cct ggc ccc tcc      843
Arg Ala Arg Gln Ala Gly Ser Ile Ala Ser Glu Val Pro Gly Pro Ser
        -1  1                5                   10

agg gag cat gat ggg cct gaa agt aac cag tgt tcc ctc cac cct ttt      891
Arg Glu His Asp Gly Pro Glu Ser Asn Gln Cys Ser Leu His Pro Phe
    15                  20          ***     25

caa gtc agc ttc cag cag ctg ggc tgg gat cac tgg atc att gct ccc      939
Gln Val Ser Phe Gln Gln Leu Gly Trp Asp His Trp Ile Ile Ala Pro
30  ***             35                  40                  45

cat ctc tat acc cca aac tac tgt aag gga gta tgt cct cgg gta cta      987
His Leu Tyr Thr Pro Asn Tyr Cys Lys Gly Val Cys Pro Arg Val Leu
                50                  55                  60

cac tat ggt ctc aat tct ccc aat cat gcc atc atc cag aac ctt gtc     1035
His Tyr Gly Leu Asn Ser Pro Asn His Ala Ile Ile Gln Asn Leu Val
            65                  70                  75

agt gag ctg gtg gat cag aat gtc cct cag cct tcc tgt gtc cct tat     1083
Ser Glu Leu Val Asp Gln Asn Val Pro Gln Pro Ser Cys Val Pro Tyr
        80                  85                  90

                    [S2]
aag tat gtt ccc att agc atc ctt ctg att gag gca aat ggg agt atc     1131
Lys Tyr Val Pro Ile Ser Ile Leu Leu Ile Glu Ala Asn Gly Ser Ile
    95                  100                 105

ttg tac aag gag tat gag ggt atg att gcc cag tcc tgc aca tgc agg     1179
Leu Tyr Lys Glu Tyr Glu Gly Met Ile Ala Gln Ser Cys Thr Cys Arg
110                 115                 120                 125

tga cggcaaaggtgca
STOP
```

FIGURE 6: Alignment of GDF9 and GDF9B protein sequence with other members of the TGFβ superfamily members for which structures have been determined

The furin processing site is indicated as a solid gray block at the start if the sequences. The mature processed protein begins at amino acid residue position 4. Conserved cystein molecules involved in disulphide bonds are shown in grey shading. Numbers along the bottom provide a relative reference to amino acid position, but do not represent the real amino acid residue number of each protein because gaps have been introduce to allow alignment of conserved protein regions. The asterisk * indicates the conserved systeine that is present in all other TGFβ family members except GDF9 and GDF9B, and which is responsible for the interchain disulphide bond present in most dimers. Boxed letters indicate the [787] serine (S) in GDF9 which is changed to phenylalanine in the mutants (position 86 on this diagram), and the [S2] serine (S) in GDF9B which is changed to isoleucine (position 118 in this diagram). Bold letters above the sequences refer to the conserved histidine involved in dimer hydrogen binding (H) in BMP7 and TGFβ3, and the conserved serine and leucine residues required for receptor binding (S, L) in BMP2.

```
BMP2 Human        REKRQAKHKQRKRLKS------------------------SCKRHPLYVDFS-DVGWND
BMP7/OP1 Human    RSIRSTGSKQRSQNRSKTPKNQEALRMANVAENSSSDQRQACKKHELYVSFR-DLGWQD
GDF9 SHEEP        RHRRDQESASSELKKPLVPASVNLSEYFKQFLFP----QNECELHDFRLSFS-QLKWDN
GDF9 Human        RHRRGQETVSSELKKPLGPASFNLSEYFRQFLLP----QNECELHDFRLSFS-QLKWDN
GDF9 Mouse        RRRRGQKAIRSEAKGPLLTASFNLSEYFKQFLFP----QNECELHDFRLSFS-QLKWDN
GDF9B SHEEP       RRARQAGSIASEVPGPSREHDGPE--------------SNQCSLHPFQVSFQ-QLGWDH
GDF9B Mouse       RSVRQACSIESDASCPSQEHDGSV--------------NNQCSLHPYKVSFH-QLGWDH
GDF9B Human       RRTRQADGISAEVTASSSKHSGPE--------------NNQCSLHPFQISFR-QLGWDH
TGFB2 Human       RKKRALDAAYCFRNVQDN----------------------CCLRPLYIDFKRDLGWK-
TGFB3 Human       RKKRALDTNYCFRNLEEN----------------------CCVRPLYIDFRQDLGWK-
TGFB1 Human       RHRRALDTNYCFSSTEKN----------------------CCVRQLYIDFRKDLGWK-
                  0         10        20        30        40        50

                                              H                           S
BMP2 Human        WIVAPPGYHAFYCHGECPFPLADHLNSTNHAIVQTLVNSVN-SKIPKACCVPTELSAISM
BMP7/OP1 Human    WIIAPEGYAAYYCEGECAFPLNSYMNATNHAIVQTLVHFINPETVPKPCCAPTQLNAISV
GDF9 SHEEP        WIVAPHKYNPRYCKGDCPRAVGHRYGSPVHTMVQNIIHEKLDSSVPRPSCVPAKYSPLSV
GDF9 Human        WIVAPHRYNPRYCKGDCPRAVGHRYGSPVHTMVQNIIYEKLDSSVPRPSCVPAKYSPLSV
GDF9 Mouse        WIVAPHRYNPRYCKGDCPRAVRHRYGSPVHTMVQNIIYEKLDPSVPRPSCVPGKYSPLSV
GDF9B SHEEP       WIIAPHLYTPNYCKGVCPRVLHYGLNSPNHAIIQNLVSELVDQNVPQPSCVPYKYVPISI
GDF9B Mouse       WIIAPRLYTPNYCKGICTRVLPYGLNSPNHAIIQSLVNELVNHSVPQPSCVPYNFLPMSI
GDF9B Human       WIIAPPFYTPNYCKGTCLRVLRDGLNSPNHAIIQNLINQLVDQSVPRPSCVPYKYVPISV
TGFB2 Human       WIHEPKGYNANFCAGACPYLWS---SDTQHSRVLSLYNTINPEASASPCCVSQDLEPLTI
TGFB3 Human       WVHEPKGYYANFCSGPCPYLRS---ADTTHSTVLGLYNTLNPEASASPCCVPQDLEPLTI
TGFB1 Human       WIHEPKGYHANFCLGPCPYIWS---LDTQYSKVLALYNQHNPGASAAPCCVPQALEPLPI
                  60        70        80        90        100      * 110

                  L
BMP2 Human        LYLDENEKVVLKNYQDMVVEGCGCR
BMP7/OP1 Human    LYFDDSSNVILKKYRNMVVRACGCH
GDF9 SHEEP        LAIEPDGSIAYKEYEDMIATKCTCR
GDF9 Human        LTIEPDGSIAYKEYEDMIATKCTCR
GDF9 Mouse        LTIEPDGSIAYKEYEDMIATRCTCR
GDF9B SHEEP       LLIEANGSILYKEYEGMIAQSCTCR
GDF9B Mouse       LLIETNGSILYKEYEGMIAQSCTCR
GDF9B Human       LMIEANGSILYKEYEGMIAESCTCR
TGFB2 Human       LYYIGK-TPKIEQLSNMIVKSCKCS
TGFB3 Human       LYYVGR-TPKVEQLSNMVVKSCKCS
TGFB1 Human       VYYVGR-KPKVEQLSNMIVRSCKCS
                  120       130       140
```

Figure 7. Examples of the pattern of progesterone concentrations in plasma of actively immunized ewes. Antigen used for immunization and the ewe identification numbers are shown at the top of each graph. Markings by vasectomized rams are indicated with arrows. Day 0 = corresponds to the beginning of thrice weekly sampling period.

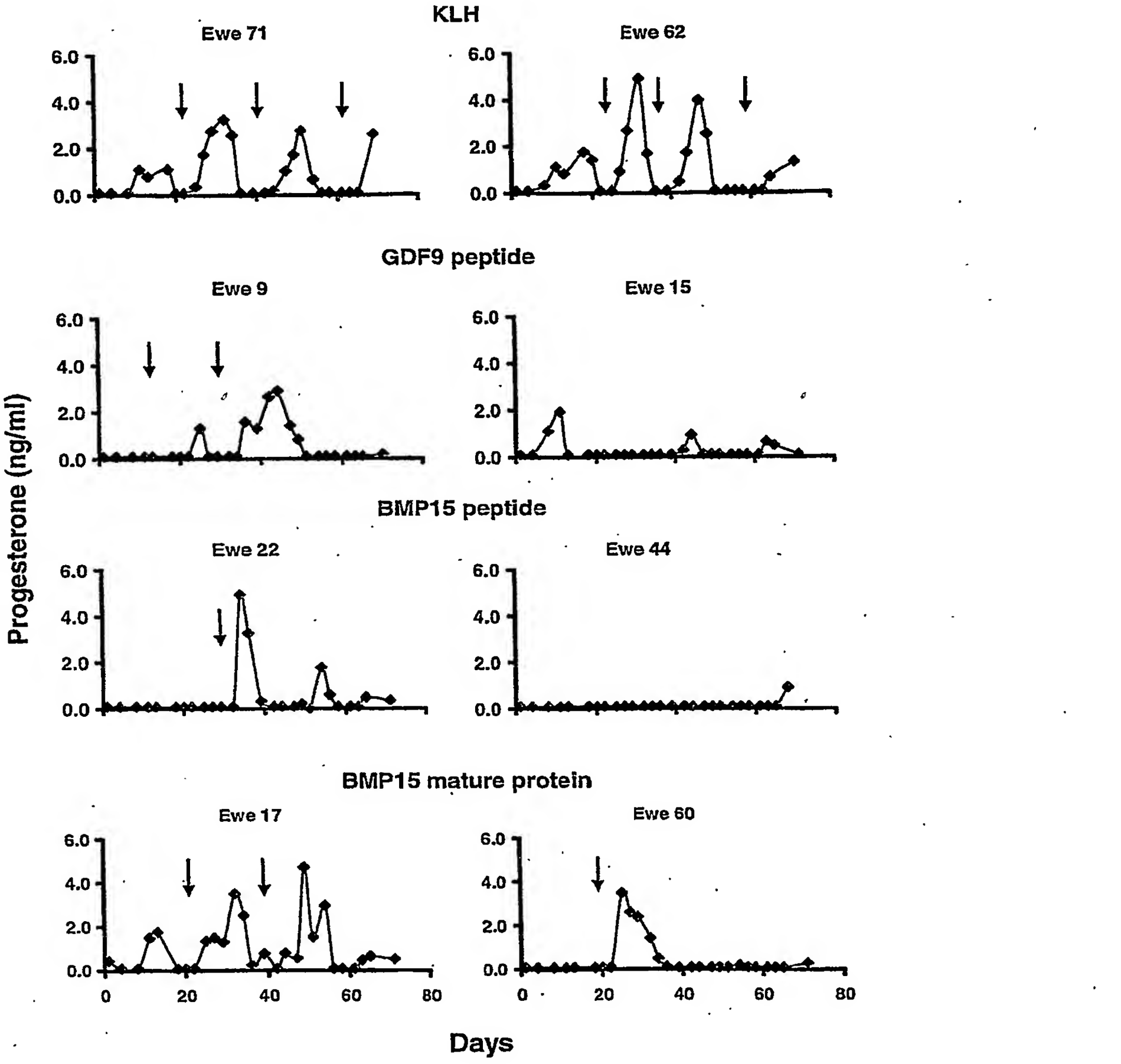


Figure 8. Average concentrations of progesterone in plasma following synchronization of luteal regression. Ewes were administered 100 ml of KLH, GDF9 peptide or BMP15 peptide antiplasma i.v. 4 days before synchronization with Estrumate (i.e. $PGF_{2\alpha}$, arrowed).